27

JUL 1 1933

Vol. 30, No. 6 PSYCHOLOGICAL REVIEW PUBLICATIONS June, 1933

Psychological Bulletin

EDITED BY

EDWARD S. ROBINSON, YALE UNIVERSITY

HOWARD C. WARREN, PRINCETON UNIVERSITY *(Review)*
SAMUEL W. FERNBERGER, UNIV. OF PENNSYLVANIA *(J. Exper. Psychol.)*
WALTER S. HUNTER, CLARK UNIVERSITY *(Index)*
HERBERT S. LANGFELD, PRINCETON UNIV. *(Monographs)*

WITH THE CO-OPERATION OF

G. W. ALLPORT, HARVARD UNIVERSITY; J. E. ANDERSON, UNIVERSITY OF MINNESOTA; C. M. DISERENS, UNIVERSITY OF CINCINNATI; W. T. HERON, UNIVERSITY OF MINNESOTA; J. A. McGEOCH, UNIVERSITY OF MISSOURI; K. S. LASHLEY, UNIVERSITY OF CHICAGO; J. T. METCALF, UNIVERSITY OF VERMONT; R. PINTNER, COLUMBIA UNIVERSITY.

LEARNING NUMBER

Edited by

JOHN A. McGEOCH

CONTENTS

General Reviews and Summaries:
Adult Learning: FLOYD L. RUCH, 387. *Sleep and Retention:* EDWARD B. VAN ORMER, 415.

Book Reviews: 440.

Communications: 459.

PUBLISHED MONTHLY (EXCEPT AUGUST AND SEPTEMBER)
FOR THE AMERICAN PSYCHOLOGICAL ASSOCIATION
BY THE PSYCHOLOGICAL REVIEW COMPANY
PRINCETON, N. J.

Entered as second-class matter at the post-office at Princeton, N. J., with an additional entry at Albany, N. Y.

Publications of the American Psychological Association

EDITED BY

HOWARD C. WARREN, Princeton University (*Review*)
S. W. FERNBERGER, University of Pennsylvania (*J. Exper. Psychol.*)
WALTER S. HUNTER, Clark University (*Index and Abstracts*)
HENRY T. MOORE, Skidmore College (*J. Abn. and Soc. Psychol.*)
HERBERT S. LANGFELD, Princeton University (*Monographs*)
EDWARD S. ROBINSON, Yale University (*Bulletin*)

HERBERT S. LANGFELD, Business Editor

PSYCHOLOGICAL REVIEW

containing original contributions only, appears bi-monthly, January, March, May, July, September, and November, the six numbers comprising a volume of about 540 pages.

PSYCHOLOGICAL BULLETIN

containing critical reviews of books and articles, psychological news and notes, university notices, and announcements, appears monthly (10 numbers), the annual volume comprising about 720 pages. Special issues of the Bulletin consist of general reviews of recent work in some department of psychology.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

containing original contributions of an experimental character, appears bi-monthly, February, April, June, August, October, and December, the six numbers comprising a volume of about 900 pages (from Jan. 1, 1933).

PSYCHOLOGICAL INDEX

is a compendious bibliography of books, monographs, and articles upon psychological and cognate topics that have appeared during the year. The Index is issued annually in June, and may be subscribed for in connection with the periodicals above, or purchased separately.

PSYCHOLOGICAL ABSTRACTS

appears monthly, the twelve numbers and an index supplement making a volume of about 700 pages. The journal is devoted to the publication of non-critical abstracts of the world's literature in psychology and closely related subjects.

PSYCHOLOGICAL MONOGRAPHS

consist of longer researches or treatises or collections of laboratory studies which it is important to publish promptly and as units. The price of single numbers varies according to their size. The Monographs appear at irregular intervals and are gathered into volumes of about 500 pages.

JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY

appears quarterly, April, July, October, January, the four numbers comprising a volume of 448 pages. The journal contains original contributions in the field of abnormal and social psychology, reviews, notes and news.

ANNUAL SUBSCRIPTION RATES

Review: $5.50 (Foreign, $5.75). **Index:** $4.00 per volume.
Journal: $7.00 (Foreign, $7.25). **Monographs:** $6.00 per volume (Foreign, $6.30).
Bulletin: $6.00 (Foreign, $6.25). **Abstracts:** $6.00 (Foreign, $6.25).
Abnormal and Social: $5.00 (Foreign, $5.25). Single copies $1.50.
Current numbers: Journal, $1.25; **Review,** $1.00; **Abstracts,** 75c; **Bulletin,** 60c.

COMBINATION RATES

Review and Bulletin: $10.00 (Foreign, $10.50).
Review and J. Exper.: $11.00 (Foreign, $11.50).
Bulletin and J. Exper.: $12.00 (Foreign, $12.50).
Review, Bulletin, and J. Exper.: $16.00 (Foreign, $16.75).
Review, Bulletin, J. Exper., and Index: $19.00 (Foreign, $19.75).

Subscriptions, orders, and business communications should be sent to the

PSYCHOLOGICAL REVIEW COMPANY

PRINCETON, N. J.

The Psychology of Human Learning: A Bibliography

By JOHN A. McGEOCH
University of Missouri

A bibliography of experimental and theoretical papers upon the problem of human learning, including memory.

1,200 titles of books and papers, covering the period from the work of Ebbinghaus to and including 1930.

A limited number of paper covered reprints of this bibliography, from the *Psychological Bulletin*, January, 1933, are available at fifty cents per copy.

62 pages 50 cents (postpaid)

▲▲▲

PSYCHOLOGICAL REVIEW COMPANY
PRINCETON, NEW JERSEY

Psychological Review Publications

Original contributions and discussions intended for the Psychological Review should be addressed to

Professor Howard C. Warren, Editor PSYCHOLOGICAL REVIEW,
Princeton University, Princeton, N. J.

Original contributions and discussions intended for the Journal of Experimental Psychology should be addressed to

Professor S. W. Fernberger,
Editor JOURNAL OF EXPERIMENTAL PSYCHOLOGY,
University of Pennsylvania, Philadelphia, Pa.

Contributions intended for the Psychological Monographs should be addressed to

Professor Herbert S. Langfeld, Editor PSYCHOLOGICAL MONOGRAPHS,
Princeton University, Princeton, N. J.

Reviews of books and articles intended for the Psychological Bulletin, announcements and notes of current interest, and *books offered for review* should be sent to

Professor E. S. Robinson, Editor PSYCHOLOGICAL BULLETIN,
Institute of Human Relations, Yale University, New Haven, Conn.

Titles and reprints intended for the Psychological Index should be sent to

Professor Walter S. Hunter, Editor PSYCHOLOGICAL INDEX.
Clark University, Worcester, Mass.

All business communications should be addressed to

Psychological Review Company,
Princeton, New Jersey.

PSYCHOLOGICAL MONOGRAPHS

AN EXPERIMENTAL INVESTIGATION OF EYE MOVEMENTS IN LEARNING TO SPELL WORDS

By Luther C. Gilbert

Mr. Gilbert has made an intensive study of the rôle of eye movements in learning to spell. He finds that "the teaching of spelling in the early years should involve a definite method of instruction." Teachers as well as experimental psychologists will find much valuable information in this monograph.

No. 196 viii + 81 pp. $1.50

PSYCHOLOGICAL REVIEW COMPANY

PRINCETON, N. J.

Vol. 30, No. 6 June, 1933

THE
PSYCHOLOGICAL BULLETIN

ADULT LEARNING

BY FLOYD L. RUCH
University of Illinois

In this paper the writer will attempt to survey the studies upon adult as compared with youthful learning and upon age changes in adult learning beyond the twentieth year, and to interpret their results in the light of certain methodological considerations. Age differences in the following phenomena will be considered: learning ability; form of the learning curve; retention; transfer and inhibition; and reminiscence. Learning is here used to mean objectively measured improvement resulting from a series of practice trials. According to this definition, studies of the phenomena of immediate memory such as digit span, fidelity of report, etc., are not logically to be included in this review.

I. Methodology

A critical reading of the experimental articles on adult learning reveals three important problems of methodology which are in many instances only imperfectly solved. This is to be expected when one considers the great complexity of the factors involved. In certain cases, however, a more critical evaluation of the experimental variables and a more studied presentation of the data on the part of the authors would have greatly enhanced the value of their conclusions. It must be pointed out in this connection that some of the studies considered in this review have as their principal aim the study of some other problem, and consequently present only incidental observations or data on the subject of age differences. The three most frequently encountered methodological problems are: (1) to obtain a random sampling of individuals, thus insuring that they vary systematically in age only; (2) to obtain the same degree of motiva-

tion throughout the entire age range; and (3) to present the data in such a manner as to control the variable factors of initial and final performance in the learning tasks at different age levels.

Selection. The problem of selecting a representative population is always difficult, and the additional problem of describing briefly the selective factors operating in choosing a group of subjects is almost insoluble. The reviewer must be content to trust the judgment of the author who, as the only person in direct contact with the many complex variables operating, is in a position to estimate or control the effect of those factors. In certain instances, however, it is possible to detect definite selective factors revealed by the data themselves or by the author's description of the method of obtaining subjects. For example, certain studies have employed children in the elementary and high schools. This method involves systematic factors. The brighter children complete their school work at a younger age than the duller ones; hence a school population is curtailed at the upper end of the distribution. There is another selective factor at work also. The duller children who have not yet reached the age where school attendance is no longer compulsory remain in school; but after that age has been reached many of them drop out to go to work, thus introducing a curtailment at the lower end of the distribution. The amount to which these two variables counterbalance one another depends in part, quite obviously, upon the age limit of compulsory school attendance and the severity of its enforcement, which in turn varies with the community.

Certain selective factors operate in the case of college populations. Younger college students are more intelligent than older ones who spent more time in completing their preparatory work. Thus Garrett (13), to choose a single example, finds a correlation of —.11 between age and Thorndike score in a population of 158 men at Columbia College averaging 19.4 years of age with an S.D. of 2.0 years, of whom 40 per cent were Sophomores, 47 per cent Juniors, and 13 per cent Seniors. However, the data of Jones, Conrad, and Horn (21) on a more representative population show no marked decline in Army Alpha test performance over a corresponding age range. It follows from the above considerations that learning tasks correlating highly with intelligence test performance would be expected to give spuriously high negative correlations with age on college populations. This type of criticism will be specifically applied to certain of the studies to be reviewed.

Motivation. The problem of motivation is even more trying than

that of selection. Common sense leads one to believe that older people are in general less motivated than younger ones. Thorndike (**36**) holds that this is only natural since many people cease learning upon reaching adulthood because they feel that they have learned all they need to know. Even the casual observer is convinced that the aged individual frequently "lets-down" or gives up the active struggle for knowledge, being content to rest upon the achievements of a busy youth. Another problem closely related to that of motivation is the loss through disuse of the habit of learning. Of two groups of adult learners of equal initial ability, it is only reasonable to expect that those who have resumed active efforts to learn after a long period of inactivity would succeed less well than another group composed of individuals who have been learning consistently throughout their entire adulthood. It is equally reasonable to expect that the degree of "mental rustiness" might vary with the task. We might expect the athletic person to learn a new game more easily than the person who has given no time to sports or, conversely, the adult scholar to learn verbal material more efficiently than the individual who has been during his youth engaged in non-intellectual pursuits.

Initial and Final Ability. The comparison of young and old learners with regard to absolute learning ability is rendered peculiarly difficult by virtue of the fact that initial ability often varies with the age group and the task. Thorndike (**36**) has employed two methods in attempting to eliminate this factor: (1) the method of common points of mastery; and (2) the method of partialing out initial ability. The first method consists in plotting the learning curves of the individual subjects and indicating some common point of mastery in the early part of the trial series and, further along in the trial series, another point which is also common to all of the compared individuals. In this manner it is possible to indicate the rate of learning of each subject in terms of the number of trials required to pass from one common point to another. This method can be used only when the curves of the subjects do exhibit common points which are far enough apart to give a series of intervening trials long enough to be reliable. The ultimate validity of this method waits upon the demonstration that the form of the learning curve and the rate of its acceleration depend uniquely upon the conditions of learning imposed during the experiment and are quite independent of the previous conditions operating to bring the learning of the subject up to the initial common point of mastery. For example, suppose that a given task has been learned in a hit or miss fashion through uncontrolled

practice in everyday life to a certain point of mastery which represents the ability of the subject at the outset of the formal experiment. Now suppose that learning has progressed from a zero point up to the before-mentioned common point under the same experimental routine as was imposed during the later learning. Will the form of the learning curve from the common point on be identical under these two conditions? The complete validity of the method of common initial and final points of mastery depends upon an affirmative answer to this question. Before this method can be accepted as a standard control of differential initial and final ability of compared individuals, it must be experimentally shown that previous learning and forgetting of a task do not influence the form of the learning curve between the two common points of mastery. In the meantime the method of common points of mastery appears to be more defensible than methods involving the expression of rate of improvement as the percentage that the gain is of the initial score when initial ability is not constant for all subjects. Furthermore, the expression of gain as a percentage in the manner just described is faulty because it assumes that the arbitrary units yielded by the experiment correspond to true units, a condition which would occur only by the sheerest chance; and because it makes conflicting and unwarranted assumptions regarding the true location of the zero point and the form of the learning curve.

The method of common points of mastery suffers from the statistical limitation that individual performance fluctuates from one trial to another although the general trend of the learning curve is upward. To eliminate this source of unreliability it would be necessary to smooth each individual curve, a procedure which involves an enormous amount of labor and implies, if arithmetic means are to be used, that the units are equal at all ranges. A variation in this method is possible when individual scores are available. This variant of the method of common points involves the use of initial and final points of mastery which are averages (preferably medians) of the scores by trial of all of the individuals in each of the compared groups. Since the units of performance in the usual learning experiment are arbitrary and probably vary in value at different ranges of ability, the median should be selected as the average since it is based on the relative positions of the curves only and does not consider the exact magnitude of the measures entering into it. In applying the method of average initial and final points to published work it is usually necessary to work with curves based on mean scores by trial since most workers present their data in that form. The method of

average initial and final common points is at best a makeshift device to be applied by inspection when considering the significance of learning curves in the absence of the original data upon which they are based. Underlying the method of common average points are certain assumptions beyond those implied by the individual score variant of the method of common points, the validity of which can be considered only in the light of empirical evidence. It suffices to indicate that if the various learning curves entering into the determination of the common points are of the same general form and if the units of practice are not too unequal at different ranges of ability, the average common point method should rank a number of differing groups in the same order as they would have when ranked according to the average of the individual common point learning ability scores. Besides this, the common average point method does not supply individual scores for correlation work. It must be pointed out that neither of these methods permits the direct comparison of the same individual on different tasks. Nor has it been demonstrated that either of these methods yields a measure of "true learning ability" independent of past experience, although it appears that certain arguments could be presented to support this contention, particularly in the case of the individual score variant of the method.

The use of the method of partial correlation in eliminating the effects of different initial and final abilities of compared subjects is open to most of the interpretive hazards of the partial correlation technique in general, and, since it also requires individual scores, is subject to any limitations arising out of the fact that learning experiments do not yield true unit scores. Frequently the nature of the material learned or the conditions of the experiment are such as to necessitate the use of this technique to the exclusion of any other as, for example, in experiments upon the learning of school subjects where daily measurement of progress is impossible and where the nature of the material learned changes from time to time.

The careful study of the published work on the problem of age differences in learning ability reveals a considerable need for more careful control of the factor of initial ability. Specific instances will be discussed in which an adequate control of this factor would have greatly enhanced the value of the obtained results. However, to limit this review to those experiments which are entirely above criticism in this regard would mean the omission of the majority of the studies on original learning and many of the studies on other phenomena.

II. Learning Ability

Nonsense Materials. Four authors report work on nonsense materials showing the relationship between increases in age beyond the twentieth year and a composite performance influenced by both initial ability and rate of learning working together in undetermined amounts.

H. L. Hollingworth (18) performed a brief experiment on transcribing a code from a key with 534 adults ranging in age from a little below twenty to forty-five years. The total population was cast into five age groups which were practically alike in four tests of intelligence. The average score in speed of transcribing was 68 for the subjects below twenty years of age, 64 for those aged twenty to twenty-four years, 60 for those aged twenty-five to twenty-nine years, 58 for those aged thirty to thirty-four years, and 57 for those aged thirty-five years or more. Since the author does not describe the conditions of learning, it is impossible to determine the extent to which these age differences depend upon differences in initial ability. Willoughby (44) gives growth curves based on about 600 subjects for abilities in eleven tests including the digit-symbol test from Army Beta. The conditions of testing were such that learning could take place although the total scores for the subjects combine both initial ability and speed of learning. Willoughby's older subjects were the parents or grandparents of his younger ones and are therefore to be regarded as differing essentially from them in age only. It is possible that a small factor operates to reduce the ability of the older subjects since they were related and might therefore be expected to show some regression as compared with the younger group, the younger group having been selected first and having been distinctly superior to the race norm. There is also the possibility of another selective factor. Not all of the parents or grandparents of the younger children were willing to coöperate in the experiment. It is pertinent to inquire as to the possible relationship between ability and willingness to serve as a subject in the experiment. Although the effect of correlation between young and old subjects could be ignored as being small, the selective factor of willingness or unwillingness to coöperate must be regarded as uncontrolled in this experiment. Performance in this test increases in a fairly rectilinear fashion from about seven to eighteen years of age, declines steadily and gently to the age of fifty when the performance is about equal to that of the twelve-year-olds. Beyond the age of fifty there is a more rapid decline so that by the time the age of fifty-seven years is reached the average performance

is about equal to that of the ten-year-olds. In a later paper (**45**) the same author reports the findings of an incidental study made on 300 of his original group of subjects. Upon completing the transcribing the subjects were asked to write as many of the number combinations as they could remember. The age curve for this performance follows the same general course as that of the substitution itself.

Fisher (**10**) gave a digit-symbol test to 382 subjects aged nine to thirty years. Fifty of these were in their twenties. The average number of correct substitutions in 10 thirty-second periods of practice was 112 for the nine-year-olds, 141 for the ten-year-olds, 180 for the eleven-year-olds, 207 for the twelve-year-olds, 228 for the thirteen-year-olds, 222 for the fourteen-year-olds, 215 for the fifteen-year-olds, 221 for the sixteen-year-olds, and 245 for the adults (twenty to thirty years old). Since the critical ages of between sixteen and twenty are not represented, it is impossible to determine exactly the point of prime for this ability. The dip in the growth curve coming at the fourteenth year indicates that some sort of selective factor must have been operating. One should hesitate to label this a "pubescent dip" until it is quite clear that selection does not account for it. Unfortunately the age of puberty comes at about the time when the brighter children are leaving the elementary schools for the high schools and at an age when the duller students are piling up in the eighth grade. Another selective factor enters here since the duller students, having finished the conventional eighth grade schooling, tend to drop out on one pretext or another to go to work. The curves based on Willoughby's data are not subject to such selective influences and do not show marked irregularities at any age.

Thorndike (**36**) reports a number of extremely interesting experiments including some on what he terms "sheer modifiability" or nonsense learning. In one experiment the subjects were told to transcribe a series of disconnected words in terms of a code in which each letter of the alphabet stood for some other letter. The experiment was performed by graduate students of education divided into six different groups but working under uniform conditions. There were eight work periods, each of three minutes, separated by intervals of from 80 to 45 seconds spent in counting and recording the number of letters transcribed or in resting. Learning curves are given for the three age groups: 28 subjects aged twenty to twenty-four years; 139 aged twenty-five to thirty-four years; and 104 aged thirty-five years or over. The average number of letters correctly transcribed in the first three minutes was 59.8 for the younger group, 56.0 for

the middle-aged group, and 47.3 for the older group; while the average gain (final minus initial scores) based on the first and last work periods was 28.0 (P.E. 1.2) for the younger group, 27.9 (P.E. 0.7) for the middle-aged group, and 22.8 (P.E. 0.7) for the older group. A check experiment in which 21 individuals were continued for eight additional practice periods gives the same order of rank for the age groups. At the end of the main experiment the subjects were asked to write each letter of the alphabet with its code equivalent. It should be remembered that nothing was said in the original directions about learning the code, the entire emphasis being upon the task of transcribing the letters as rapidly as possible. The average results are as follows: the young group learned 10.4 letters, the middle group learned 8.3 letters, and the old group learned 6.3 letters. Since the older group transcribed fewer letters than the younger group, it is fairer to express learning in terms of the number of repetitions of the code elements rather than in total time spent transcribing. In terms of total time spent the older group remembered only 66 per cent as many letters as the younger group, but in terms of total repetitions the older subjects did 77 per cent as well. Whichever method may be regarded the fairer may be disputable, but the fact remains that the old group is distinctly less efficient than the young group.

In another experiment on nonsense learning 150 women, students in summer school, tried to associate three-letter nonsense syllables with three-place numbers during eighteen exposures of the paired associates. At the end of the exposure period each student filled in the numbers on a blank giving the first member of the pair. Two forms of the test were given, each form containing twenty-five pairs. The two forms given on different days show a corrected reliability of .80. The results based on the averages of the two forms are as follows: 76 subjects aged eighteen to twenty-four years learned 21.9 of the 25 syllables; 42 subjects aged twenty-five to thirty-four years learned 18.2 syllables; and 32 subjects aged thirty-five to fifty-four years learned 14.0 syllables. Here again we observe a marked decrease with age in immediate memory for nonsense material.

In a similar experiment on adult prisoners at Sing Sing, aged seventeen to fifty-four years, Thorndike found ability to learn nonsense material to decline with advancing age. Starting with zero knowledge of letter-number code, fifty-four subjects transcribed a list of 420 words into numbers, for eight practice periods of three minutes each. The correlation between age and rate of transcribing during the final three-minute period is —.23 for a group

of fifty-four subjects. Knowledge of the code was tested by having the subjects write the letters of the alphabet in scrambled order and put next to each letter the number corresponding to it. The correlation between age and knowledge of code at the end of the eight practice periods was —.07 (n is 59). Since all subjects started with zero knowledge of the code, this coefficient may be considered more representative than the one for rate of transcribing.

The results of the series of studies just reviewed, based on nearly 2,000 cases in all, are in complete agreement that the composite performance scores, combining initial ability and increases due to learning, decline slowly with age between twenty and fifty or sixty years of age. The value of the experimental results would have been greatly enhanced had all the authors employed some technique to permit the expression of gain as distinct from a composite of gain and learning.

Sensory-Motor Learning. Thorndike (**36**) applied the method of common stages of mastery to Hollingworth's (**17**) data for the learning of a simple motor coördination. In order to obtain points common to all subjects it was necessary to break the population up into two groups. Both groups showed the same age range, but one contained nine and the other six subjects. The correlations of age with rate of gain between common points was —.79 for the larger and —.895 for the smaller group. The correlation of age with the approximate physiological limit was even closer (—.79 for the larger and —.92 for the smaller group).

Lashley (**23**) studied the rate of improvement of a group of men and boys ranging in age from fourteen to thirty-six years in learning archery. Gain was defined as the difference between the average distance the first 40 of a series of 300 shots missed the bull's eye and a similar measure for the last 40 shots. Lashley gives curves showing the relationship between amount of gain and chronological age to be very slight. His data are not presented in such a way as to permit an estimation of the effect of initial ability on this relationship. All subjects were naïve in that they had had no previous formal training in archery, but it is to be expected that their initial ability did vary with age.

Noble (**27**) studied the improvement of three thirty-two-year-old and three twenty-year old men in throwing basketball goals. The older group gained 28.3 per cent while the younger group gained 25.8 per cent. The slightly better learning of the older group is explained by the author as being due to their greater height. The method of presenting the data does not permit the application of the

technique of common average points or of any other method of estimating the influence of differential initial ability.

Johnson (20) reports a study in which a group of 76 college freshmen were taught to walk a tight wire. Each trial consisted in starting at one end and walking until the subject lost his balance and fell from the wire. The criterion of mastery adopted was the number of trials required to learn to walk from one end of the wire to another without falling off. Under these conditions the amounts of practice afforded by the trials are very unequal, but this inequality does not, perhaps, greatly influence the reliability of the total score since the tendency would be for short and long trials to figure in the total score in proportions determined by the skill of the learner. This skill was found to correlate highly with intelligence, the Pearson coefficient of correlation between intelligence and number of trials required to meet the criterion being —.69. The correlation between age and number of trials necessary to reach the criterion was .34. This correlation seems unusually high for so short an age range and is probably attributable to the lower intelligence of the older subjects.

Thorndike (36) describes certain experiments on sensory-motor learning. In one of these, adults were taught to write with the left hand. The subjects were cast into three groups on the basis of age, each of the three age groups showing about equal average intelligence as measured by the CAVD tests. Quality was measured on the Thorndike scale for judging handwriting, and speed was measured in terms of the number of letters written per minute. A group of 17 young subjects aged twenty to twenty-five years gained 1.11 points in quality and 35 letters per minute in speed, and an older group of 17 subjects aged thirty-five to fifty-seven years gained 1.16 points in quality and 18 letters per minute in speed over a period of 15 hours of distributed practice. A middle group of five subjects aged twenty-six to thirty-four years gained .96 points in quality and 32 letters per minute in speed. In another experiment a group of five individuals aged twenty-three to thirty-eight years practiced typewriting. Their progress was about as fast as that of younger students in commercial high school. In an effort to measure what he has termed sheer modifiability independent of past acquisition the following experiment was devised: The subjects were asked to draw blindfolded, lines 3, 4, 5, and 6 inches long. The subject first drew 600 lines without knowledge of the correctness of the lengths of his lines. Then on each of seven days (consecutive except for Sundays and an occasional absence) 600 lines were drawn with the information of "right" or "wrong" given after each line. "Right" meant

that the line drawn was within one-eighth inch of the standard for the 3-inch line or within one-fourth inch for the others. Two age groups were employed: a young group of twelve individuals aged sixteen to twenty years with an average CAVD score of 412; and an old group of the same size aged thirty-three to forty-eight years with an average CAVD score of 419. Before training, the performances of the two groups were equal. The curves of the two groups tend to diverge with practice in favor of the younger group, the older group making but 64 per cent as much gain as the younger.

Kitson (**22**) made a study of the rate of improvement of a group of 40 men aged eighteen to forty-nine years in hand-composition typesetting under a special wage stimulus. Thorndike (**36**) shows that the correlation between chronological age and improvement is .036 when length of experience and initial ability are partialed out. This study is especially interesting since it indicates the possible rôle of high motivational level in reducing the effects of advancing age upon learning ability.

Freeman (**12**) reports some unpublished work of Gonnelly in which an eight-year-old child and its two parents were put through a practice series of tossing shot into a tumbler 1,800 times. The average absolute gain for the parents about equals that for the child, although the curves for the two parents start and finish at a higher point. The original data being unavailable, the reviewer has attempted to apply by inspection the method of common points to the curves given by Freeman. The composite curve of the two parents starts at about 32 successful throws in the first series of 100 attempts and rises to about 86 at the 18th series; while that of the child starts at about 6 and rises to about 46 successful throws in series 18. When the two common points of 32 and 46 are selected it is seen that the child required about 600 trials to make this gain as against about 425 for the mid-parent. Thus it is seen that the child learns this task with considerably greater difficulty than its parents.

Tachibana (**35**) in an article reviewed in English by Yoshioka (**35**) reports experiments in which young and old learners were compared on learning to pick up dry beans with chop sticks, tapping, and mirror drawing. There were five old and three young subjects. The experiments involved daily practice for from two to four weeks. The main practice was with the right hand but tests with the left hand were inserted at regular intervals. In the chop-stick test a woman of eighty-nine was compared with another of twenty-one years. The older woman showed no improvement while the younger gained considerably. In the tapping test a sixty-year-old woman and a sixty-

nine-year-old man showed no improvement for either hand. Two controls, an eighteen-year-old and a twenty-year-old girl, showed typical learning curves for both hands. In the mirror drawing experiment two old subjects were used. A seventy-year-old man drew a five-pointed star seen in a mirror five times daily for ten days and reduced his time to one-third of the initial score. The other subject, who did only one day's series of ten trials, showed a slight improvement. The slight improvement of the two aged subjects in learning to trace a star in the mirror might be due to their mastery of the intellectual features involved rather than to an actual increase in motor coördination. This interpretation is consistent with the observed lack of gain in tapping and in picking up beans with chop sticks, tasks which had possibly been learned up to the physiological limit of motor coödination as determined by age. However, according to this explanation one would expect to find some cross transfer from the right to the left hand, since any central factor would be common to both. No such transfer was observed.

Renshaw (30) studied the effect of practice on cutaneous localizing movement in four children aged eight to eleven years and seven adults aged eighteen to sixty-five years. The subject located on the volar forearm and hand a point stimulated by experimenter. No knowledge of results was given. Large practice effects were noted in the case of the children and smaller effects in the adults. The practice curves of the old and young group for localization on the forearm do not overlap, but in the case of the data for localization on the hand the method of common average points of mastery can be applied to the data presented by the author. The children pass from an average error of 10.5 mm. to one of 8 mm. in about 15 attempts; while the adults require about 90 attempts to make this same amount of improvement.

These studies on sensory-motor learning indicate that increases in age beyond twenty years have an adverse effect upon motor learning. The only study failing to agree is that of Noble, who has himself pointed out the probable source of error. If one were to conclude solely from those studies in which a fairly complete control of initial and final ability is possible, the conclusions would be the same. The three studies comparing the learning of young children with that of adults do not permit a definite conclusion as the total number of subjects is too small to permit the drawing of growth curves.

Meaningful Verbal Materials. Thorndike (36) has worked over the original data of Hollingworth (17) in an attempt to determine the effect of age on learning ability. Fifteen subjects divided into

three groups with age ranges of nineteen to twenty-four, twenty-four to twenty-six, and thirty-three to thirty-nine years were compared on learning ability in naming opposites, mental calculation, and motor coördination. In the naming of opposites the rate of improvement was expressed for each subject as the amount of response time required to improve from the rate of 250 opposites named in 450 seconds to that of 450 opposites named in the same length of time. The median times required to make this amount of gain were: 195 seconds for the youngest group; 170 seconds for the middle-aged group; and 180 seconds for the oldest group. Thorndike gives the correlation of —.12 between age and rate of gain for these data. The data from the mental calculation test (adding mentally a given series of numbers to 17) were treated in a similar manner except that here the method of measuring gain was to determine the number of repetitions required to improve from a performance level common to nearly all of the subjects to a higher common point of mastery. The average number of repetitions required to improve from a rate of 320 seconds per series of 50 mental additions to 160 seconds per series was 12.2 for the young, 11.3 for the middle, and 13.1 for the old group. The correlation of age with gain was —.20.

Worcester (**46**) studied the relationship between age and ability to memorize prose with thirteen summer session students ranging in age from twenty-one to fifty-eight years. The materials were presented visually in one part of the experiment and auditorially in another part. The reviewer has calculated coefficients showing the relationship between age and time required to learn to be —.13 for the auditorially presented prose and —.35 for the visually presented prose.

Thorndike (**36**) presents the results of a series of experiments on the learning of Esperanto. This material was chosen because the learning of it was felt to be representative of the learning of any body of organized knowledge. The main experiments were with university students ranging in age from twenty to fifty-seven years. Twenty hours were spent in studying, ten in class and ten at home. All subjects had the same teaching. Progress was measured by four tests including vocabulary, printed directions, oral directions, and paragraph reading. The subjects were cast into three groups showing almost equal average intelligence test scores and initial ability. The youngest group containing 18 subjects aged twenty to twenty-five years gained 31.5 points, the 9 subjects aged twenty-six to thirty-four years gained 26.3 points, and the 21 subjects aged thirty-five to fifty-seven years gained 24.7 points. The superiority of the younger

adults was shown almost entirely on the oral directions test. Results from other experiments not reported in detail indicate that ability to learn Esperanto falls off about 22 per cent between the ages of twenty-two and forty-two years, according to Thorndike's summary of the data.

Sorenson (34) reports an experiment on adult learning in which the achievement of three groups of adults in education courses is analyzed by means of the partial correlation technique. Both experimental groups contained elementary school teachers who were candidates for the Bachelor's Degree in Education. Group I was composed of 58 individuals who had had no immediately previous course work. Groups II and III containing respectively 41 and 54 subjects had had immediately previous course work. Four measures were made on all of the subjects: chronological age; the number of hours spent in study as reported by the subjects; the scores on the Miller Analogies Test; and the achievement of each subject as measured by an objective examination on the materials of the course. The partial correlation between age and achievement independent of hours of study and analogies score was —.32 for the group having no immediately previous course work. For the other two groups the corresponding coefficients were .07 and —.12. These results seem to indicate that practice in learning may be an important factor in maintaining the ability to learn in spite of the advance of age. Since this is a very important point, it would have been better had the author included a test of initial knowledge of the materials of the course. One might expect the group having had immediately previous course work on related topics to show higher initial knowledge and, as a consequence, less variability in amount of gain. This study is outstanding as the only one which attacks directly the problem of the possible rôle of practice in learning in retarding the decline of learning ability.

Two studies have been made comparing adult learners with children below twenty. Joseph Peterson (29) gave his rational learning test to nineteen individuals of various ages. The younger group of subjects contained one high school and ten college students, while the older group was composed of two graduate students and six business and professional adults. He found much overlapping between the two groups, although the older did distinctly better, requiring but 6.0 as against 10.0 trials to master the problem. Three of the younger group did worse than the poorest of the older subjects, a woman fifty years of age. It should be pointed out that, while all subjects started with zero knowledge of the letter-digit combinations, it is possible that

age differences in knowledge of technique of attack did exist. Since the exact ages of the subjects are not given, it is impossible to compute a correlation coefficient showing the relationship between age and learning efficiency. Chen (5), in an article abstracted in English by Chow (5), gives the results of an experiment in which a modified form of Peterson's test was administered to his mother and to two of his nephews (their ages are not stated in the review). The two nephews learned the test much faster than the mother.

With a group of 158 undergraduate men, Garrett (13) found the correlations between age and a composite score based on initial ability and learning efficiency to be —.19 for a digit-symbol test, —.22 for Turkish-English vocabulary, and —.25 for a modified form of the Andersonville Prison Code. The effect of partialing out intelligence as measured by the Thorndike Test is to reduce these coefficients to —.18, —.19, —.23, respectively.

There has been enough work on the adult learning of meaningful materials to open up a series of problems but not enough to permit definite conclusions. The weight of the evidence is that age beyond the twentieth year affects this sort of learning less adversely than motor learning or the learning of nonsense materials.

Maze Studies. Hicks and Carr (16) compared four adult graduate students with five children aged eight to thirteen years on learning a simple maze. The subjects were blindfolded. In seventeen trials the children made 120 errors and the adults but 54. The children also took more time and traveled more surplus distance than the adults.

Gould and Perrin (14) conducted a similar study employing a pencil maze of more complex pattern and a larger number of subjects. The fifteen children employed were, all but two, siblings of college students and had an average age of 11.6 years. These children were compared with a group of twenty-five subjects, all but one of whom were college undergraduates, with an estimated average age of twenty-one years. It is reasonable to suppose that the two groups were fairly comparable except for age. The young group made more errors and traveled more excess distance but took slightly less time than the old group.

Husband (19) compared a group of fifteen older subjects (mainly graduate students) with a group of several hundred undergraduates on learning mazes. Since all of the subjects did not learn the same maze, their performance was expressed in terms of percentile ranks based on the younger population. The average percentile ranks of the older subjects are 42 for errors, 38 for trials,

and 27 for time. The greater deficit of the older subjects on time scores may indicate a greater caution as well as a lower learning ability. Husband feels that his older subjects were fully as intelligent as his younger ones. This and the fact that the older group was actively engaged in mental work make their poorer performance seem all the more indicative of a lowered learning ability consequent upon age.

In the study of maze learning the problems presented by differences in initial and final ability are especially troublesome because of their concealment. The traditional method of scoring maze learning implies that all subjects start with zero knowledge of the maze. While it can be granted that all subjects are ignorant of the particular pattern, there is still the argument that some come to the problem better prepared by knowledge of mazes in general and that this knowledge can vary systematically with the age of the learner, thus introducing an uncontrolled variable. The fact that the first trial, unless it can be assumed to start from zero, reflects both initial ability and progress in different quantities makes it unjustifiable to employ total error or trial scores as a basis for comparing groups. It seems reasonable to suppose that previous experience would manifest itself more in the earlier than the later trials. It seems preferable, therefore, that the method of common points should be employed here also, selecting as the first common point a degree of mastery exhibited by all subjects at the end of a few trials. The best treatment of the error scores can accordingly be made, it would seem, by determining the number of errors or the number of trials required to improve from one point to another, selecting some rigid criterion of mastery as the final common point. Time scores, since they never reach zero, should be expressed for purposes of comparison as the amount of time required to improve from a lower to higher speed of traversing the maze. It must be pointed out that error scores on mazes and problems involving the learning of lists of words, nonsense syllables, lines of poetry, lists of paired associates, and the like, can not be compared with improvements in time scores in performance on motor tasks, mazes, or other problems, since the former types of learning have a logical limit whereas the latter have a physiological limit. This is a point which is frequently neglected in published studies on learning.

The method of common points can not be applied to the results of the Hicks and Carr and the Gould and Perrin studies since individual scores by trial are not given. The former study, however, gives average scores by successive trials, from which it appears that the

differences between children and adults would be less when rate of learning is determined by the method of common average points.

III. Form of the Learning Curve

Four workers present data showing the influence of age on the form of the learning curve. In the experiment of Hicks and Carr reviewed in the previous section the children as compared with the adult learners made a larger proportion of their total errors during the early trials. When time scores and surplus distance were employed as measures of learning efficiency, similar results were obtained, the children as compared with the adults showing a larger proportion of total errors or time during the earlier trials. The average performance of the children was more variable than that of the adults. These results taken together indicate less insight on the part of the children at the outset of the experiment. The experimental conditions of the Gould and Perrin study were outlined in the previous section. When the error, time, and excess distance scores for trials one to twenty are plotted in terms of the percentage that each trial score is of the total score, the curves for the children and the college students tend to have the same form and show much overlapping. During the first part of the trial series the adults make slightly fewer errors, take less time, and travel less excess distance; but after the tenth trial this relationship becomes reversed. The results for the children are more variable than those for adults. The conclusions from this study are in essential agreement with those drawn by Hicks and Carr. Neither of the two studies reports any evidence of qualitative differences in the mode of attack between the young and old learners.

Snoddy (**33**) has extensively studied the form of the learning curve in mirror drawing as dependent upon the age of the learner. His subjects ranged in age from young children to individuals eighty years of age. He was able to distinguish two phases in the learning curve for mirror drawing: (1) adaptation; and (2) facilitation. The adaptation phase is characterized by a rapid increase in efficiency as a result of practice. This rapidity of learning is explained by Snoddy as being due to the growth of neural processes perhaps in the manner implied in the theory of neurobiotaxis of Kappers. The facilitation phase is characterized by a slow regular increment in efficiency of performance and is held by Snoddy to result from condensation of the previously established neural patterns. Adults and more especially senescent individuals give learning curves which are relatively lacking in the adaptation phase and resemble the facilitation

phase of the learning curves of young subjects. The young learners show at the outset an adaptation phase followed by facilitation. These views are consistent with the current biological thought which regards childhood as a time of easy and rapid growth and senescence as a period of great stability. Snoddy is the only worker among those whose papers have been reviewed who attempts to explain age differences in the form of the learning curve on a definitely neurological basis.

Thorndike (36) in comparing the form of the learning curves for his Teacher's College group of subjects in code transcription eliminates the effect of differing amounts of total gain by expressing the gain for each practice period as a percentage of the total gain. The curves thus plotted for the three age-groups are practically identical except for chance fluctuations. In the experiment on learning to draw lines of a given length, the curves for the two groups differ in that the younger learners continue to improve after the older subjects have almost reached their physiological limit. The learning curve of the group aged thirty-five to fifty-seven years approaches in form that of the sixty to eighty-year-old subjects of Snoddy, while the curve for Thorndike's younger group, aged twenty to twenty-four years, closely resembles the adaptation curve described by Snoddy. It appears that Snoddy has stated a fundamental descriptive principle which holds true for at least two types of motor material even though a complete acceptance of the underlying neurological explanation be held in abeyance.

IV. Retention

Henderson (15) made a study of the retention of verbal materials over periods of two days and four weeks in children ten to sixteen years old and in university summer-session students. To eliminate the effects of differing amounts of original learning, Thorndike (36) paired each of the sixteen adults with pupils of varying ages who showed the same degree of original learning. The loss in mastery of the material at the end of a month's time was 7.1 for the adults as against 7.4 for children aged fourteen to sixteen years; 7.0 for adults as against 5.7 for children aged twelve and thirteen years; and 8.1 for adults as against 4.1 for children aged ten and eleven years. Thus it appears that the retentivity of school children for meaningful material becomes equal or inferior to that of adults at about year thirteen or fourteen. Similar comparisons show graduate students to lose an average of 12.8 items as against 11.9 for college under-

graduates. University students show an average loss of 12.6 as against 13.0 for high school students.

Lyon (**24**) studied the learning and relearning of series of digits, nonsense syllables, and words, and of prose passages and poetry in subject groups whose modal ages ranged from fourteen to thirty-four years. The conditions under which the original learning was accomplished make it impossible to estimate the effects of differential initial ability, on the one hand, or of differing amounts of overlearning, on the other. When his data for the different age groups are compared by stating the amount of time saved in relearning as a percentage of the original learning time, no significant differences are noted.

Downey and Anderson (**7, 8**) found that a woman forty-one years of age relearned the previously acquired ability to write while reading silently or aloud in less time than was required by a twenty-one-year-old man.

Achilles (**1**) investigated the phenomena of recall and recognition in groups of widely differing ages. She found no consistent differences between graduate and undergraduate students in the recall or recognition of words, geometric forms, or proverbs, but her results indicate that the younger students excel the older ones in recognition of nonsense syllables.

Chen (**5**), in the study reviewed elsewhere, found that his mother retained the numbers of the Peterson rational learning test after a year's time, whereas the two young nephews had forgotten several of the numbers.

Worcester (**46**) found evidence with thirteen summer-session students aged twenty to fifty-eight years that increased age within these limits is accompanied by a loss of retentivity. The conditions of original learning have already been described. Tests of recall were made at intervals of one, two, and seven days following the original learning. The reviewer has computed correlations between age and the amount of material retained at the end of one and seven days. These are —.32 and —.42, respectively, for the auditorially presented prose, and —.32 and —.33 for the visually presented prose.

Williams (**43**), in a study of the phenomena of reminiscence, found that partially learned lists of unconnected words were retained equally well by four groups of subjects averaging about ten, thirteen, sixteen, and twenty-one years of age. The ten-year-old group learned 42 per cent of 30 words, the thirteen-year-old group learned 46 per cent of 40 words, the sixteen-year-old group learned 48 per cent of 50 words, and the adult group learned 54 per cent of 50 words. Taking the amount originally learned as 100 per cent, curves of

retention were plotted for all four of the groups. These curves show no reminiscence in any of the groups and are identical in form and rate of decline except for minor fluctuations attributable to chance.

Thorndike (36) presents correlation coefficients showing that retention correlates positively with age in adults. Prisoners at Sing Sing Prison transcribed a list of 420 words into numbers according to a code. The code was printed in large characters on the blackboard and exposed only during the practice periods. The men transcribed for eight three-minute periods, stopping at the end of each period to count the letters transcribed. Just before the practice series the code was exposed, used for a sample word, and recovered. The letters were then dictated in scrambled order with instructions to give their code equivalents. These papers were collected and the test repeated. The results of these two tests showed practically zero knowledge of the code at the outset. Consequently, the number of letters correctly given when the same test was repeated at the end of the learning period can be taken as a measure of improvement in knowledge. After a period of one week without practice the same test was again given as a measure of retention of code-knowledge. The men were then practiced for four three-minute periods as before. The scores of these periods were used as a measure of rate following rest. The measure of retention in rate of substitution is the extent to which the subjects took up their work after the rest minus the rate before the rest. A high degree of retention is thus indicated by a negative score of low magnitude and a low degree of retention is indicated by a negative score of greater magnitude. The correlation between age and retention for 54 subjects aged seventeen to fifty-four years was practically zero (.06). Since the older subjects were of lower intelligence than the younger ones as measured by the Pintner Non-Language Mental Test, and since the younger subjects showed greater rate before the rest period, Thorndike controlled these variables by means of the partial correlation technique. The partial coefficient of correlation of age with retention independent of intelligence and rate before the rest period is .13.

In determining the relationship between age and retention of code-knowledge, retention was measured as the knowledge of the code after the rest interval and not as the difference between knowledge after and before. Retention measured in this fashion is not the complement of forgetting but a composite measure combining the amount of original learning minus the amount forgotten. The correlation between age and retention of code-knowledge was .05. When the

effects of intelligence and knowledge at the end of the first day are partialed out a coefficient of .23 is obtained.

Thorndike also compared 12 subjects about twenty-two years old with 10 others twenty years older, in memory for Esperanto and in retention of ability to write with the wrong hand. The results are difficult to interpret since the initial ability of the subjects varied. In wrong-hand writing the older subjects lost .24 points and the younger group lost .16 points by the end of seventy-seven days of no practice. In number of letters written per minute the loss of the older group was 2.7 and of the younger group 5.5. When these results are stated as the percentage that the loss is of the initial gain, the direction of the difference is reversed. At the end of seventy-seven days the subjects in the Esperanto experiment were given tests similar in form to those used at the end of the practice period. Twelve subjects aged twenty-two to twenty-four years showed a loss of 0.2 points of combined speed and quality score while 10 subjects averaging twenty years older showed a loss of 4.9 combined points. The greater loss of the older subjects seems the more apparent when it is noted that their initial gain was less and that their knowledge at the end of the learning period was also less than that of the younger subjects.

The results of the various studies on retention point to no definite conclusions. Part of this lack of agreement may arise from differences in defining the experimental conditions. In the traditional retention experiment the subjects learn the material to a certain objective degree of mastery common to all of them. At the end of a period of no practice the retention of the subjects is tested in various ways. If the material permits, he recalls as many of the items as possible, or designates those previously learned, or relearns the same task to the previous degree of mastery. The first two methods do not present any particular difficulty when the degree of mastery of the subjects just before the rest period is constant. The method of relearning often employs the amount of savings as a means of expressing retention. According to this method the saving in the number of trials required to relearn the material is expressed as a percentage of the number of trials required to learn the material originally. This procedure is defensible only when the compared individuals start with equal initial ability. This ideal is closely approximated when comparable groups of the same age are subjected to varied conditions during the period of forgetting; but when subjects of different age are compared, pains must be taken to see that the initial ability of all subjects is the same. For example, suppose

that an older group of subjects starts at a greater initial ability and reaches the objective criterion of mastery in fewer trials than a younger group. The same amount of absolute forgetting will give a much higher per cent of savings score for the younger than for the older group. It seems that the safest procedure in comparing groups or individuals where variable initial ability is a factor is to practice all individuals to an initial common point of mastery, determine the amount of practice required to bring all of them up to a second common point, preferably the *logical* or *physiological* limit of the task, and, taking the amount of practice required to improve between the common points, express the savings in relearning as a percentage of the original learning score as here defined. This method is based on the assumption that differences in ability at the outset of the experiment result from differences in chance opportunities to learn the particular task before the imposition of formal experimental conditions, and from differences in native ability to learn. Although the method of partial correlation might appear to present a means of controlling the effects of differences in initial and final ability, its logic is less obvious than that just presented.

There are three studies which do not permit comparisons between individuals of different ages but are interesting enough to merit brief mention. Ebbinghaus (**9**) relearned stanzas of *Don Juan* after a lapse of twenty-two years. At the time of relearning no evidence of memory was introspectively revealed, nor was any objective recall apparent. However, the number of trials required to relearn the stanzas were less than that required to learn stanzas not previously studied. Warren (**40**) reports two cases of recrudescence of youthful memory after a long interval had elapsed. The author's father at ninety years of age suddenly recalled a poem learned seventy-five years previously. The other case was that of a man who recalled at the age of eighty-three years a Freshman oration learned sixty-nine years previously and not recited during the interim. Titchener (**37**) gives an introspective account of learning to mastery a poem memorized at the age of ten.

V. Transfer and Inhibition

Meumann (**25**) held that the concept of a general memory must include certain general functions which coöperate in all memorial acquisition in addition to the mere acquisition of a particular content, and also a sharing by all related special memory groups of profit gained from the training of any one of them. He reported evidence

indicating that adults are less susceptible than children to improvement in memorizing ability as a result of practice in memorizing. Three adults aged thirty-six to fifty-four years did not show so much improvement in memory as a result of practice as a group of university students.

Thorndike (**36**) continued his experiment on code transcribing with prisoners to permit a study of transfer. Fifty-four subjects were given eight three-minute practice periods in transcribing followed by a seven-day period of no practice. There were then given four three-minute periods of relearning followed by an equal amount of time spent in reverse substitution with the same code. Two days later the men practiced a new code in which the letter-digit combinations were changed to confuse. In reverse substitution, learning in which previous associations are helpful, the correlation between age and transfer was .01 or —.02 when intelligence and first-day final ability were held constant by the partial correlation technique. Transfer was measured as the rate in reverse substitution minus the rate before the reversal. The correlation between age and transfer to the disarranged code learning where previously acquired associations interfere was —.03, or —.17 when intelligence and retained rate are partialed out. Transfer was here measured as the rate of using the new code minus the retained rate of using the first code. So small a difference in correlations based on only 54 cases would hardly justify Thorndike's conclusion that younger men are better able than older men to prevent or overcome hindrance by learning which transfers negatively.

Sleight (**32**) noted incidentally in connection with his study of memory and formal training that younger individuals show more pronounced transfer effects than adult learners, but the fact that the same materials and methods were not used in the two age groups prevents any quantitative statement of the results. Dallenbach (**6**) practiced a group of third-grade children on visual apprehension materials and found positive transfer. Whipple (**41**) and Foster (**11**) had previously employed similar materials with competent adults and found no transfer. Van Wayenburg (**39**) found positive transfer between two motor tasks in children aged thirteen and fourteen years but not in a group of adults. One is tempted to conclude from these experiments that increases in age below adulthood are accompanied by decreases in amount of transfer. It should be pointed out, however, that children starting at a lower point of mastery than adults

UNIVERSITY OF ILLINOIS LIBRARY

have greater opportunities to show transfer, the adults being limited by their closeness to their physiological limit.

Brown (3) investigated the formation of various habits of card sorting involving the overcoming of interfering habits. In one experiment with 26 college students it was found that the older subjects were more adversely affected than younger ones by interfering habits. This he explains as being partly due to the lower intelligence of the older subjects. In a second experiment on 18 subjects differently selected this result was not verified.

Rugg (31) conducted a study of the transfer effect of a course in descriptive geometry on ability to employ spatial concepts and symbols. The subjects were 413 students at the University of Illinois who ranged in age from seventeen to twenty-six years. Tests measuring the ability in question were given to the experimental and control groups before and after the training period. When the dividing line is drawn at age twenty, 60.2 per cent of the younger group show a gain as against 62.5 per cent of the older group. When the subjects below eighteen years of age were compared with those aged twenty-two years or more the results are reversed, 64.4 per cent of the younger subjects showing gain as against 58.8 per cent of the older ones. This study fails to reveal any consistent age difference within the limits studied.

Whitely (42) conducted a series of experiments in which he studied the effect on learning and recall of arousing an apperceptive mass. The results indicate that such prior activities as were calculated to arouse the apperceptive mass actually impeded learning and retention. This effect does not vary with age in groups of grammar school, high school, and college subjects.

VI. Reminiscence

Ballard (2) made a study of phenomena of memory which he called reminiscence. In certain of his experiments young children, adolescents, and young adults were given copies of poetry to read for fifteen minutes. At the end of this period they were instructed to write as much of the poem as they could remember. This immediate recall was taken as 100 per cent and the amount of subsequent recall was expressed as a percentage of it. Nothing was said about further tests of memory. Sub-groups were tested after intervals of one, two, etc., days. Groups averaging about six years of age showed progressive increases up to the third day when they averaged 160 per

cent of its immediate recall. Of the subjects aged about twelve years one sub-group showed a maximum recall of 110 per cent and this on the second day. The oldest subjects showed no evidence of reminiscence, all of the sub-groups giving less than 100 per cent of their immediate recall. The results of Ballard can not, however, be interpreted as showing a decrease in reminiscence consequent to increasing age because his older subjects learned more initially and thus had less opportunity to show reminiscence. Partial learning is a necessary condition of this phenomenon.

Norsworthy (**28**) studied the learning and recall of English equivalents of German words in a group of 73 college students. The students were tested twice with a period of one month intervening. Eight subjects earned the same score on both occasions, 26 students did better on the retest, and 39 of them did worse. Brown (**4**) conducted a short experiment in which a list of 48 words was read to a class of 193 students. The words were first read according to the Berrol memory system thus: "Hat; He had a hat on his head; Hat." The list was then repeated without the mnemonic device. The average immediate recall was 25.48 words and the average recall after a half hour spent listening to a class lecture was 26.77. Although small, this difference indicates some reminiscence. Nicolai (**26**) presented series of 10, 20, and 30 objects to eights adults and ten children about twelve years of age. Both groups recalled about the same number of objects immediately after. Three tests for reminiscence were made at intervals of 1, 24, and 96 hours. The children recalled more objects later than immediately on all tests; while the adults recalled more objects later than immediately on four out of nine tests.

Williams (**43**) conducted an extensive experiment on reminiscence in which he attempted to insure the same partial degree of initial learning at all ages by giving the older subjects more material to learn. Poetry and disconnected words were used but only the former showed reminiscence. More than 2,000 subjects divided into age-groups averaging about ten, thirteen, sixteen, and twenty-one years of age studied lines of poetry for five minutes. During this period the ten-year-old group learned 25 per cent of 24 lines; the thirteen-year-old group learned 27 per cent of 32 lines; the sixteen-year-old group learned 36 per cent of 40 lines; and the adult group learned 38 per cent of 40 lines. Thus it is seen that the length of the task was not increased enough to bring about the same proportionate amount of original learning in the four age groups, although the

UNIVERSITY OF ILLINOIS LIBRARIES

condition of partial learning was obtained. The two younger groups showed reminiscence but the two older ones did not. It seems doubtful if the slightly greater amount of original learning in the older groups is enough to account for or invalidate the results. A complete control of the factor of original learning could best be obtained by practicing each subject to the same point of partial mastery of the same amount of material. Aside from the considerations just presented, it seems definitely established that there is a progressive decrease with age in amount of reminiscence.

BIBLIOGRAPHY

1. Achilles, E. M., Experimental Studies in Recall and Recognition. *Arch. of Psychol.*, 1920, **6**, No. 44.
2. Ballard, P. B., Obliviscence and Reminiscence. *Brit. J. Psychol. Monog.*, 1913, **1**, No. 2.
3. Brown, W., Habit Interference in Sorting Cards. *Univ. of Calif. Pub. in Psychol.*, 1914, No. 53.
4. Brown, W., To What Extent Is Memory Measured by a Single Recall? *J. Exper. Psychol.*, 1923, **6**, pp. 377–382.
5. Chen, H. C., Capacity to Learn of an Aged Person. *Chin. J. Psychol.*, 1923, **2**, No. 1. (Reviewed: Chow, *Psychol. Abs.*, **2**.)
6. Dallenbach, K. M., The Effect of Practice upon Visual Apprehension. *J. Educ. Psychol.*, 1914, **5**, pp. 387–404.
7. Downey, J. E., and Anderson, J. E., Automatic Writing. *Amer. J. Psychol.*, 1915, **26**, pp. 161–195.
8. Downey, J. E., and Anderson, J. E., Retention of Skill After Lapse of Practice: Simultaneous Reading and Writing. *Amer. J. Psychol.*, 1917, **28**, pp. 396–408.
9. Ebbinghaus, H., *Grundzüge der Psychologie*, 1905.
10. Fisher, V. E., A Few Notes on Age and Sex Differences in Mechanical Learning. *J. Educ. Psychol.*, 1927, **18**, pp. 562–564.
11. Foster, W. S., The Effect of Practice upon Visualization and upon Reproduction of Visual Impressions. *J. Educ. Psychol.*, 1911, **2**, pp. 11–21.
12. Freeman, F. N., *How Children Learn*, 1917.
13. Garrett, H. E., The Relation of Tests of Memory and Learning to Each Other and to General Intelligence in a Highly Selected Adult Group. *J. Educ. Psychol.*, 1928, **19**, pp. 601–613.
14. Gould, M., and Perrin, F. A. C., A Comparison of the Factors Involved in the Maze Learning of Human Adults and Children. *J. Exper. Psychol.*, 1916, **1**, pp. 122–154.
15. Henderson, E. N., A Study of Memory for Connected Thoughts. *Psychol. Rev. Monog.*, 1903, **5**, No. 23.
16. Hicks, V. C., and Carr, H. A., Human Reactions in a Maze. *J. Animal Behav.*, 1912, **2**, pp. 98–125.
17. Hollingworth, H. L., The Influence of Caffein on Mental and Motor Efficiency. *Arch. of Psychol.*, 1912, **3**, No. 22.

18. HOLLINGWORTH, H. L., *Mental Growth and Decline*, 1927.
19. HUSBAND, R. W., Certain Age Effects on Maze Performance. *J. Genet. Psychol.*, 1930, **37**, pp. 325–328.
20. JOHNSON, G. B., Study in Learning to Walk the Tight Wire. *Ped. Sem.*, 1927, **34**, pp. 118–128.
21. JONES, H. E., CONRAD, H., and HORN, A., Psychological Studies of Motion Pictures. II. Observation and Recall as a Function of Age. *Univ. Calif. Pub. Psychol.*, 1928, **3**, pp. 225–243.
22. KITSON, H. D., A Study of the Output of Workers under a Particular Wage-Incentive. *Univ. J. Business*, 1922, **1**, pp. 54–68.
23. LASHLEY, K. S., The Acquisition of Skill in Archery. *Papers from the Tortugas Lab. of the Carnegie Institution of Washington*, 1915, **7**, pp. 105–128.
24. LYON, D. O., The Relation of Quickness of Learning to Retentiveness. *Arch. of Psychol.*, 1916, No. 34.
25. MEUMANN, E., *The Psychology of Learning*, 1913. (Trans., Baird.)
26. NICOLAI, F., Experimentelle Untersuchungen über das Haften von Gesichtseindrücken und dessen zeitlichen Verlauf. *Arch. f. d. ges. Psychol.*, 1921, **42**, pp. 132–149.
27. NOBLE, S. G., The Acquisition of Skill in the Throwing of Basket Goals. *School and Society*, 1922, **16**, pp. 640–644.
28. NORSWORTHY, N., Acquisition as Related to Learning. *J. Educ. Psychol.*, 1912, **3**, pp. 214–218.
29. PETERSON, JOSEPH, Experiments in Rational Learning. *Psychol. Rev.*, 1918, **25**, pp. 443–467.
30. RENSHAW, S., The Errors of Cutaneous Localization and the Effect of Practice on the Localizing Movement in Children and Adults. *J. Genet. Psychol.*, 1930, **38**, pp. 223–238.
31. RUGG, H. O., *The Experimental Determination of Mental Discipline in School Studies*, 1916.
32. SLEIGHT, W. G., Memory and Formal Training. *Brit. J. Psychol.*, 1911, **4**, pp. 386–457.
33. SNODDY, G. S., Learning and Stability. *J. Appl. Psychol.*, 1926, **10**, pp. 1–36.
34. SORENSON, H., Adult Ages as a Factor in Learning. *J. Educ. Psychol.*, 1930, **21**, pp. 451–459.
35. TACHIBANA, K., On Learning Process of the Aged. *Jap. J. Psychol.*, 1927, **2**, pp. 635–653. (Reviewed: Yoshioka, *Psychol. Abs.*, **2**.)
36. THORNDIKE, E. L., *Adult Learning*, 1928.
37. TITCHENER, E. B., Relearning After Forty-Six Years. *Amer. J. Psychol.*, 1923, **34**, pp. 468–469.
39. VAN WAYENBURG, —., *Observations on Manual Dexterity*. VII International Cong. Psychol., 1923.
40. WARREN, H. C., Two Cases of Long Latent Memory. *Psychol. Bull.*, 1918, **15**, pp. 207–209.
41. WHIPPLE, G. M., The Effects of Practice upon the Range of Visual Attention and of Visual Apprehension. *J. Educ. Psychol.*, 1910, **1**, pp. 249–262.
42. WHITELY, P., Dependence of Learning and Recall upon Prior Intellectual Activities. *J. Exper. Psychol.*, 1927, **10**, pp. 489–508.

UNIVERSITY OF ILLINOIS LIBRARY

43. WILLIAMS, O., A Study of the Phenomenon of Reminiscence. *J. Exper. Psychol.*, 1926, **9**, pp. 368–387.
44. WILLOUGHBY, R. R., Family Similarities in Mental-Test Abilities. *Genet. Psychol. Mon.*, 1927, **2**, pp. 235–277.
45. WILLOUGHBY, R. R., Incidental Learning. *J. Educ. Psychol.*, 1929, **20**, pp. 671–682.
46. WORCESTER, D. A., Memory by Visual and Auditory Presentation. *J. Educ. Psychol.*, 1925, **16**, pp. 18–27.

SLEEP AND RETENTION

BY EDWARD B. VAN ORMER

St. Joseph's College for Women and Wagner College

The purpose of this paper is the review and evaluation of the experimental literature which bears upon the topic of sleep and retention.[1] The first to suggest the possibility that retention may be affected by sleep was Ebbinghaus (6). In his study of retention as a function of time, he comments upon the discrepant results at the 8.8-hour and 24-hour periods. Table I below shows part of his data

TABLE I

PERCENTAGES RETAINED AFTER CERTAIN TIME INTERVALS ACCORDING TO VARIOUS INVESTIGATORS USING SAVINGS METHOD

(from van Ormer, Arch. Psychol., 1932, # 137)

Time in hrs.	Ebbinghaus Obtained	Ebbinghaus Theoretical	Ebbinghaus Diff.	Radosavljevich	Finkenbinder	Luh	Boreas	van Ormer
1 hr.	44.2	46.7	−2.5	70.7	72.8	65.9	64.19	44.1
2 hrs.	....	....	....	....	69.4	....	.....	38.5
4 "	....	....	....	....	66.4	54.9	.....	34.1
8–9 "	35.8 (8.8 hrs.)	34.5	+1.3	47.4 (8 hrs.)	65.5 (8 hrs.)	....	57.59 (9 hrs.)	23.9
16 "	....	....	....	....	63.0	....	.....	....
24 "	33.7	30.4	+3.3	68.9	57.8	52.1	59.21	....
36 "	....	....	....	....	58.8	....	.....	....
48 "	27.8	28.1	−0.3	60.9	55.5	47.7	50.40	....

compared with the expected or theoretical results computed from his equation for the curve.[2] Regarding this discrepancy we can say that the loss for the 15-hour period from 8.8 hours to 24 hours is 2.1 per cent and the loss for the 24-hour period from 24 to 48 hours is 5.9 per cent, or almost three times as much. Regarding this greater loss he says, " Such a condition is not credible, since in the case of all the

[1] The writer is indebted to Professor K. M. Dallenbach for valuable criticisms of the manuscript; to Professor J. A. McGeoch for the references to Foucault (8), to Hunter (10), to Robinson (21), and to Spight (24); to Professor R. S. Woodworth for the reference to and translation of Boreas' article (2).

[2] Three of his calculations differ from the theoretical by one-half of one per cent or less; two by 1 to 1½ per cent; and one differs by 2½ per cent; the 24 hour one by 3.3 per cent, offering the greatest discrepancy. Of these divergences, four are plus (better retention) and three are minus (6, p. 106).

other numbers the decrease in the after-effect is greatly retarded by an increase in time. It does not become credible even under the plausible assumption that night and sleep, which form a greater part of the 15 hours but a smaller part of the 24, retard considerably the decrease in the after-effect. Therefore it must be assumed that one of these three values is greatly affected by accidental influences. It would fit in well with the other observations to consider the number 33.7 per cent for the relearning after 24 hours as somewhat too large and to suppose that with a more accurate repetition of the tests it would be 1 to 2 units smaller. However, it is upheld by observations to be stated presently, so that I am in doubt about it" (p. 104–105; translation by Ruger and Bussenius, p. 77). These control tests to which he refers were performed later with 12 syllables instead of 13 and with the criterion of learning being one correct repetition, rather than two correct repetitions in succession.[3] The results from these tests gave approximately the same value for the 24-hour period as he obtained in the main experiment, the values being 33.7 per cent and 33.4 per cent respectively, for the main and control experiments. Apparently here is the first time that the suggestion that sleep is beneficial to retention occurs in the psychological literature, although rejected by Ebbinghaus as the explanation of his discrepancy.

Radosavljevich[4] (20), in his study of the curve of forgetting, found a decided drop at the 8-hour interval and an almost complete recovery at 24 hours, with comparatively little loss from 24 to 48 hours, the 48-hour value remaining above the 8-hour value. This result was obtained in all his data, with children and adults, and with meaningful and meaningless material; although the percentage of loss and recovery varied somewhat among the different situations. Radosavljevich thinks that the difference at 8 and 24 hours is chiefly conditioned by sleep, one being fatigued after 8 hours of waking activity and at the 24-hour period having recovered through sleep. He states later on that the reports of his *S*s verified this supposition. Apparently he means by recovery through sleep, the recovery in learning efficiency and not a better retention as the result of sleep.

[3] van Ormer and Dallenbach (28) have pointed out that although the difference between the results of Radosavljevich and Ebbinghaus has often been attributed to a difference in the criterion of learning used, such a difference does not exist, since Ebbinghaus, in his main experiments on the curve of forgetting used *two* correct repetitions in succession, as did Radosavljevich.

[4] According to the *Psychological Register*, 1929, p. 190, he now spells his name in this manner.

It is interesting to note that his discrepancies are somewhat similar to those of Ebbinghaus at the 8.8- and 24-hour intervals. A discussion of possible explanations for Radosavljevich's discrepancies will be given later.

Bean (1), who studied retention for series of consonants and for typewriting, with times of retesting of from 1 to 28 days, discusses the discrepancies which Ebbinghaus and Radosavljevich found at the 8- to 9-hour and 24-hour periods. He thinks that a better distribution of experiments on the part of Ebbinghaus would have helped to eliminate his discrepancy, apparently meaning, from the context preceding, that Ebbinghaus should have devoted more experiments to the 8.8-hour interval and not so many to the longer intervals. In reply to this it should be noted that the P.E. of the mean of Ebbinghaus' results for the 8.8-hour interval is smaller, save for the 31-day interval, than for any of the longer intervals. Bean says of Radosavljevich's recovery between the 8- and 24-hr. intervals, "This may reasonably be attributed to the fact that the 8-hour interval brought the relearning into an unfavorable time of day, while the 24-hour interval allowed sleep to intervene and brought the time of relearning back to the original hour of the day, with all conditions, external and internal, most nearly identical with those of the first learning" (p. 42). Bean offers, however, no data to settle the question, since his first test was made after an interval of one day.

Foucault (8), in a practically unknown paper[5], discusses Ebbinghaus' discrepancy at the 8.8- and 24-hour intervals, and, without giving Ebbinghaus credit for suggesting it, offers the opinion that the lesser amount of forgetting during the 15-hour period between the 8.8- and 24-hour intervals may have been due to the predominance of sleep. He then offers the hypothesis, which antedates by at least 11 years more recent suggestions of a similar nature, that, "Forgetting appears to take place in a more active fashion during the day than during night or sleep: this signifies, without doubt, that forgetting is not the work of time, that time does not act on memories with a sort of magic power, but that that which effaces memories, are the new perceptions, the new acts of the mind, which is to say that forgetting may be an act of mental opposition, or the inhibition of images by the images or perceptions which succeed

[5] Since Foucault's article is so little known, it might be mentioned that it also reports some experiments on the relation between the length of the material and the inhibitory effect, leading to the hypothesis that the percentage forgotten is inversely proportional to the length of the list, provided other conditions are equivalent.

UNIVERSITY OF MICHIGAN LIBRARIES

them" (p. 449, translation by present writer). He feels that this is substantiated by the work of G. E. Müller on retroactive inhibition. He then reports some work of his own on retroactive inhibition, testing 4 *S*s with nonsense syllables. The results show the inhibiting effect of other memorizing following shortly after the learning of a given list, as opposed to a time interval of conversation only. This, he points out, further substantiates the above hypothesis. He adds, "Forgetting is the forcing back of memories into the subconscious, more and more removed from consciousness, and this process is produced by the psychic acts following fixation" (p. 451, present writer's translation). He says further, that if an empty or completely vacant consciousness could be produced, it is his belief that retention would be as good at the end of the state as at the beginning. Feeling that such a state is impossible to attain experimentally, he calls upon the anecdotal evidence from students who report that when they are having difficulty in learning a lesson just before retiring, they will wake up in the morning, knowing the lesson, or nearly knowing it, requiring just a little rehearsal to fix it. He believes this phenomenon results from two causes: (1) memories poorly fixated at the moment of sleep continue to develop in a subconscious fashion; (2) memories poorly fixated at night are subjected to only an insignificant inhibition by the subconscious psychic acts of sleep. They have been protected by sleep against the destructive action that waking activity exercises.

Finkenbinder (7) was quite interested in Radosavljevich's discrepancy at the 8-, 24-, and 48-hour intervals, and conducted an investigation on retention of lists of 12 nonsense syllables, using 14 *S*s and testing 11 time intervals. He attempted to take care of diurnal variation in learning performance by distributing his learning times over 5 different times of the day for every time interval measured, the times of learning extending from 8:00 A.M. to 8:00 P.M. Every point on the curve for every *S* was, however, determined by only 5, or less, sittings, distributed throughout 5 different parts of the day, and every sitting consisting of the learning of only one list. Thus it would seem that rather than taking care of diurnal variation, he has simply averaged together the results from the 5 different times of the day, which results may possess unknown differences in learning performance varying with the individual *S*.[6] It

[6] Although he reports, without giving statistical reliabilities, the average learning times of 12 *S*s out of his 14, at the 5 different times of the day, he apparently makes no use of this in correction of his data. His data seem to indicate that the diurnal variation was different for his different *S*s.

would seem from his own report that he did not have reliable measures of each time interval for each *S*.

Finkenbinder did not, however, obtain Radosavljevich's discrepant results at the 8- and 24-hour intervals. He points out, in discussing Radosavljevich's results, that the latter failed to take proper account of diurnal variations in learning performance. He had several *S*s working in the morning and several in the afternoon. This was convenient for all intervals up to 24 hours, except the 8-hour relearning interval. For this interval the afternoon workers were shifted to morning. Finkenbinder reports that he has calculated that this change lowered the learning time for the 12-syllable lists, in the case of the shifted *S*s, from 8.2 to 6.6 repetitions (or one-fifth), and increased the relearning time by one-fourth. Using the data in Radosavljevich, to which Finkenbinder refers, the present writer is unable to arrive at the same result which Finkenbinder reports.[7]

On the basis of this calculation from only part of the 8-hour experiments (see note 7) of Radosavljevich and for only the 12-syllable lists, Finkenbinder attempts to correct Radosavljevich's 8-hour values, apparently for the combined results with 8-, 12-, and 16-syllable lists. This correction would increase the retention at the 8-hour period from 47 to 66 per cent; and Finkenbinder says, smooths out Radosavljevich's curve at this point. Assuming that this correction of Finkenbinder's is valid, which we think is not the case, he has still failed to notice that there remains a recovery of approxi-

[7] Of the three *S*s who were shifted from afternoon to morning at the 8-hour interval in this part of the study, one took one *more* repetition for the original learning for the 8-hour interval than for the 1-hour interval which preceded it; the other two each took one less repetition at the 8-hour interval than at the 1-hour interval. This is far from the one-fifth decrease about which Finkenbinder writes, the decrease actually is only 3.4 per cent for the 12 syllable lists, and no more than might be expected from simple variation in performance. Other experiments on the 8-hour interval, done in the early part of the investigation, are evidently not considered by Finkenbinder; nor does he consider in any case the 8 and 16 syllable lists. In these earlier experiments the practice effect is great, however, and it is impossible to know how much of the decrease in learning time which 4 *S*s shifted from afternoon to morning show at the 8-hour experiment is due to this shift, because these *S*s show decrease in learning time both before the shift and after it, when they return to the afternoon learning for tests after 24 hours. The writer believes, after studying Radosavljevich's tables, that it is impossible to prove anything about the effect of this shift of some *S*s from afternoon learning to morning learning. For some of the experiments on the 8-hour interval, the individual results are not even reported.

mately 2.9 per cent at the 24-hour period, and he makes no effort to explain the very little loss from 24 to 48 hours (see Table I).

After some further discussion, Finkenbinder concludes that Radosavljevich's exaggerated loss at 8 hours is largely due to loss of mental efficiency at the close of the day. This conclusion is out of all consistency with Finkenbinder's own report of the diurnal variation of his *S*s. Only 4 of his 12 *S*s (the data from 2 *S*s could not be used here) found 5:00 P.M. the worst time for learning and one found it the time for best learning. Moreover, Finkenbinder's data do not bear out either Radosavljevich's or his own statements about the lapse of mental efficiency at the close of the day, or loss of efficiency due to fatigue; for 4 of his 12 *S*s found 8:00 P.M. the time of best learning, while, as above, one found 5:00 P.M. the best, and only 4 found it the worst. Of course it still remains possible that some of Radosavljevich's discrepancy at 8 hours may be due to diurnal variation; for he certainly did not control it.

Aside from the question of Finkenbinder's correction of Radosavljevich's data, the former makes another, very pertinent criticism of the latter. This is to point out the tremendous practice effect which enters into Radosavljevich's results. In the latter's main experiments, after only one to three days of practice, he begins testing the time intervals of 5 and 20 minutes and of 1 and 8 hours, in the order given. A study of his tables shows this practice effect going throughout the above experiments and on into later experiments. Radosavljevich, himself, was learning lists in later experiments in one-fourth to one-fifth the number of repetitions which he needed in the first main experiments. Other *S*s showed marked practice effects as well. Even though he attempted to eliminate this difficulty by more experiments after the *S*s had become practiced, the inclusion of the early experiments in the final averages tends to vitiate his results. The more one studies Radosavljevich's data, the less is one surprised at the variations which they show.

Finkenbinder, in spite of his attempts to avoid Radosavljevich's errors, and in spite of the fact that he does not have the same discrepancy as Radosavljevich, does obtain discrepant results. There is, for example, almost twice as much lost from 16 to 24 hours as from 8 to 16 hours; and a 3.3 per cent loss from 36 to 48 hours; while the figure for 36 hours is actually 1 per cent higher than for 24 hours. He, however, does not seem to think these discrepancies sufficiently serious to warrant any definite comment, aside from the

tabular statement of the deviation of his results from the theoretical ones calculated on the basis of his equation. When one considers his procedure of *introducing* diurnal variations as a disturbing factor, and the few tests of each *S* at each time interval, it is not surprising that such discrepancies occur.

Heine (9), working under G. E. Müller and studying retroactive inhibition, calls attention to the fact that McDougall (15) mentions that the evidence from retroactive inhibition throws some light on the assertion of many persons that what is learned just before falling to sleep is remembered with unusual exactness. (Apparently Heine did not notice Ebbinghaus' suggestion that sleep might favor retention, nor see the bearing that his and Radosavljevich's discrepant results at the 8- and 24-hour periods had on this problem.) Heine then, in connection with her other work on retroactive inhibition, conducts some side experiments on this statement of McDougall's.

Here she was concerned with the comparison of retention after 24 hours when the syllables are learned just before going to bed and when they are learned with waking intervals interpolated between the learning and sleeping. Four *S*s learned lists of 10 nonsense syllables by a modified form of the saving method,[8] every *S* learning 2 lists at a sitting. A seemingly sufficient number of experiments were conducted with every *S* to make the results reliable, although statistical reliabilities are not given. The results showed that learning just before going to bed gave an average saving after 24 hours of 47 per cent; while tests with waking periods interpolated between the learning and sleep gave an average saving of only 36 per cent after 24 hours.[9] The results of all the *S*s were in favor of sleep. Heine points out that this result is still more significant when one considers that the learning disposition at night was such that with 3 *S*s, four, three, and two times respectively, the prescribed number of repetitions occurred without an errorless reproduction taking place. Another factor which makes the difference still more striking is that with three of the *S*s the learning for the "day" tests took

[8] Heine, *op. cit.*, 228–229. She recorded the number of repetitions necessary for one errorless reproduction; but had the *S* repeat each list the same number of times. Preliminary experiments determined the number of readings so that the number for one errorless reproduction would be likely to be exceeded. The number of readings was decreased as practice entered, keeping day and night the same of course.

[9] Percentages calculated by writer from data given by Heine, 231–233.

place at 8:00 P.M. rather than in the day time, and apparently only about 3 hours elapsed between the learning and sleep.

She had 2 other *S*s learn lists of nonsense syllables by the method of paired associates, with results that seem to substantiate those with the saving method. On the average 13.1 per cent was recalled after 24 hours when sleep had followed immediately upon the learning and 8.7 per cent when the learning was followed by a day of interpolated activity.[9] Heine concludes that the difference in favor of learning just before going to bed is due to the elimination of the retroactive inhibition produced by the day activities which normally follow learning.

Nicolai (19), studying reminiscence in tests of memory for objects exposed once simultaneously, finds the tendency at times, with adults, for the number of items reported to first increase at the 24-hour interval. He thinks this substantiates Radosavljevich's recovery after sleep. The case of one *S* is cited who could report 2 more objects after 4 hours of sleep (5 hours after the original presentation) than he could 45 minutes after the presentation. Nicolai states that experience itself shows the favorable effect which sleep has upon memory. One criticism of his suggestion that sleep may explain this tendency of reminiscence to first appear at 24 hours is the fact that for children of twelve and thirteen years of age he found a tendency at times for this reminiscence to begin at 1 hour; while with six-year-olds there was very little, if any, reminiscence at any time interval studied.

Luh (14), using 8 to 10 *S*s in his comparison of retention as measured by 5 different methods, found, with the relearning method, that the loss from 24 to 48 hours was approximately 1.6 times as great as the loss from 4 to 24 hours, the values being 4.4 per cent and 2.8 per cent respectively (see Table I). The other 4 methods gave a more decided drop at these periods, with the exception of recognition, which dropped only 3.1 per cent from 24 to 48 hours. Although he discusses in general the difference between the relearning curve and the other curves, he apparently fails to comment on this strange slowing up of the relearning curve between 4 hours and 1 day. Since Luh's experiments on the comparison of the five retention methods were apparently based on not more than 5 tests per *S* per method for every time interval (and in some cases on less), and because he tested written reproduction, recognition, and rearrangement, in the order given, all with the same list at one sitting, it is questionable whether Luh has really established the specific curves for these five methods, particularly when it is discovered that every

test consisted merely in the learning of one list of syllables.[10] It is worthy of note, however, in connection with the 8–9 and 24-hour discrepancies of other investigators using the relearning method, and in connection with our problem, that for his relearning curve the forgetting was less for the 20-hour period from 4 to 24 hours, a period in which sleep had more predominance, than for the 24-hour period from 24 to 48 hours.

Jenkins and Dallenbach (13), in their study of retention after sleep and waking, call attention to Ebbinghaus' and Radosavljevich's discrepant results at the 8 to 8.8 and 24-hour intervals and mention Ebbinghaus' comment on this discrepancy and the suggestion rejected by him that the discrepancy might be due to better retention after sleep, which suggestion had apparently laid unnoticed throughout the years. They aimed very definitely at the problem of what effect intervals of 1, 2, 4, and 8 hours of sleep interpolated immediately upon the learning of one list of 10 nonsense syllables have upon the curve of retention as ordinarily obtained for the same temporal periods of waking activity. Two college senior men were used as naïve *S*s in this rigorous experiment. Retention was tested by the method of retained members, approximately 8 experiments being conducted for each *S* for each time interval of waking and 8 experiments for each *S* for each time interval of sleeping. Each experiment consisted of the learning and later recall of one list of 10 nonsense syllables.

Figure I presents the results of these investigators. They found that on the average twice as many syllables were reproduced by both *S*s after sleep as after waking. The superiority of the reproductions after sleep is shown at every time interval, and is statistically reliable.[11] Notice that the sleep curve stays practically on a level from the 2-hour interval on, while the waking curve continues to drop. The curves of the individual *S*s very nearly coincide. The difference in the number of repetitions required for learning in the

[10] He reports some additional confirmatory experiments for written reproduction, recognition, and reconstruction, but not for anticipation and *relearning*. It is not clear from Luh's report whether he really controlled diurnal variation or not. Pages 13 and 21 mention an effort to do so; but what he says raises sincere doubts as to whether he really controlled it.

[11] Taking the averages and mean variations as given by the authors, and assuming a normal curve, the writer has calculated the PE of the difference for the various comparisons of recall after sleep and waking. The least reliable D/PE diff., is 4.35 for the 2-hour interval for one of the *S*s, all the other differences being above this figure.

day time rather than at night is not significant. So the difference between sleep and waking cannot be explained as due to a difference in depth of learning or number of repetitions employed in learning.

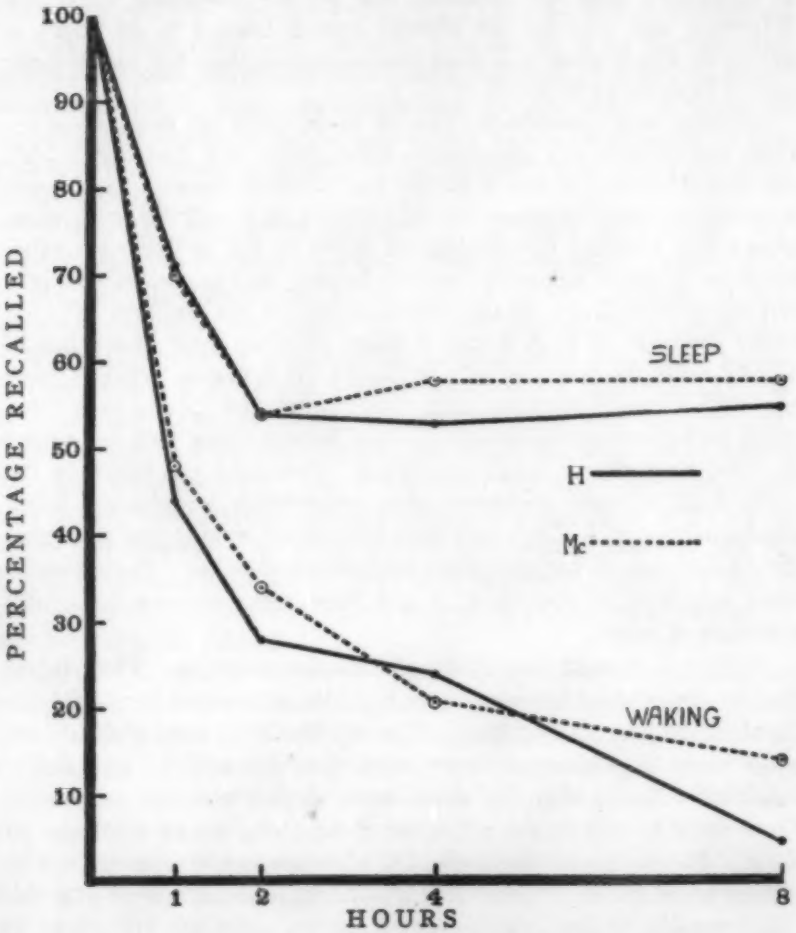

FIG. I. Average number of syllables recalled by each *S* after various time-intervals of sleep and of waking. (After Jenkins and Dallenbach, *Amer. J. Psychol.*, 1924, 35.)

Similar results in favor of sleep were obtained with an instructor who acted as *S* to test the 1- and 8-hour intervals. The intervals of 2, 4, 8, and 24 hours were also tested with one list each, using an untrained *S* who memorized the list while hypnotized and recalled it at the required time interval when put back into the hypnotic state.

At intervals of 2, 4, and 8 hours he reproduced the list perfectly, and at 24 hours reproduced 8 out of the 10 syllables. These investigators state that authorities on hypnosis disagree as to whether there is intensified or weakened memory during this state, and that of course these results with the hypnotized *S* are too few for any conclusion on the matter.[12]

Independently of Foucault (**8**), Jenkins and Dallenbach concluded that: "The results of our study as a whole indicate that forgetting is not so much a matter of the decay of old impressions and associations as it is a matter of interference, inhibition, or obliteration of the old by the new" (p. 612). They also state that: "The difference in the rate of forgetting during sleep and waking is sufficiently large to account for the variation in Ebbinghaus' investigation. His figures for the 8.8-hour, 1- and 2-day intervals are not only 'credible' but are confirmed by this study. The discrepancies in Radossawljewitsch's investigation may likewise be explained by our results" (p. 611).

The writer agrees that the discrepancy at the 24-hour interval in Ebbinghaus' curve is probably due to sleep. It is doubtful if the Radosavljevich discrepancy can be explained as due to any particular factor. Doubtless sleep (Jenkins and Dallenbach), failure to take care of diurnal variation (Bean, Finkenbinder), failure to control the practice effect (Finkenbinder), and various other factors of poor technique all contribute to give the results which he obtained.

Dahl (**5**), working under Köhler and Lewin, undertook a study of the effect of sleep on recognition, using lists of 12 nonsense syllables and series of 12 nonsense figures, and testing 14 *S*s at the time intervals of 1, 2, 4, and 8 hours of sleep and of waking activity as in the Jenkins and Dallenbach problem, which he was interested in studying by means of the recognition method. Every *S* was tested with one list of syllables and one series of figures for each time interval.

Dahl found that retention, as applied to the ability to recognize *old* material was better after the intervals of sleeping.[13] At the 1- and 2-hour intervals the ability to discriminate the *new* material was *poorer* after sleep than after waking (more false recognitions); for the 4- and 8-hour intervals the discrimination of the *new* material

[12] *Cf.*, for example, Stalnaker and Riddle (**25**).

[13] The errors given in terms of percentage of possible errors on old material, were on the average: waking 34 per cent, sleep 24 per cent, the advantage of sleep being present at all time intervals. Dahl, *op. cit.*, 296.

was somewhat *better* after sleep, than after waking.[14] The conclusion that sleep is favorable to recognition, he suggests, should be accepted with caution. After the first two hours of sleep, at least, he points out that the *S*s seemed drowsy and in an uncritical attitude, and as the results show, disposed to give more yes answers. This would favor success on the old material and errors on the new.

If we consider all the errors for a given *S*, regardless of whether they are false recognitions or failures to recognize, his results seem to show that retention as measured by recognition is better after 4 and 8 hours of sleep than after the corresponding intervals of waking, the percentage of possible errors being, on the average, for 4 hours, waking 29 per cent, sleep 23 per cent; and for 8 hours, waking 33 per cent, sleep 26 per cent. The advantage of sleep for the 1- and 2-hour intervals is not so clear, the percentage of *total* errors being slightly in favor of waking. The results are on the average: for 1 hour, waking 23 per cent, sleep 27 per cent; and for 2 hours, waking 26 per cent, sleep 29 per cent.

Dahl's results may be criticized from the standpoint that they are certainly not statistically reliable for any given *S*, since every *S* was tested only once at a given time interval, for each type of material. The reliability of his group averages cannot be ascertained, since he gives no measures of variability. We leave it with the reader to decide which is more desirable, questionable results from 14 *S*s or reliable ones from one or two? Furthermore, the reduction of the *S*s' critical attitude or the decrease in sagacity (Hollingworth) seems to be a disturbing factor in Dahl's results. He realized this and recommended that in another experiment of this sort the *S*s should be very thoroughly awake before the testing took place. In connection with this drowsiness and uncritical attitude reported by Dahl, it is interesting to note that Jenkins and Dallenbach report that *E* found it difficult to arouse the *S*s and to tell when they were awake. Often the *S*s would perform the test at night and then the next day say that they remembered nothing of it, *E* having judged by their behavior that they were awake.

Dahl offers no theory to explain his results, simply saying that if we accept the result that sleep in general favors recognition it must be assumed to be due to an influence which either sleep or

[14] The errors given in terms of the per cent of possible errors on new material, were for one hour: waking 17 per cent, sleep 31 per cent; for 2 hours: waking 15 per cent, sleep 34 per cent; for 4 hours: waking 24 per cent, sleep 20 per cent; for 8 hours: waking 30 per cent, sleep 26 per cent. Dahl, *op. cit.*, 297.

waking has upon the impressed material as in the experiments of Jenkins and Dallenbach. He says that it has not been possible to prove retroactive inhibition in the case of recognition, as it has been for other types of retention. Here he is probably referring to the work of Heine (**9**). As mentioned above (p. 9), Heine explains the result that sleep favors retention as measured by both the saving and paired associates methods, on the basis of the retroactive inhibition exercised by the day activity. Assuming this theory, Dahl argues that we should expect no advantage of sleep or of waking activity in the case of recognition, since no retroactive inhibition has been proved for this type of retention. Accordingly he does not accept retroactive inhibition as an explanation of his results and thinks that it is questionable if it can be accepted as an explanation of Heine's "sleep" experiments, or of those of Jenkins and Dallenbach.

Woodworth and Poffenberger (**30**, Chap. 10, Sect. IX), in discussing Heine's contention that retroactive inhibition does not affect recognition, conclude that retroactive inhibition does affect recognition, though probably to a less degree than it affects recall. That the effect on recognition should be smaller is to be expected, they point out, from the all-round superior efficiency of recognition over recall. They mention that Strong (**26**) found that recognition is less efficient for a large number of items than for a smaller number, and that this certainly seems to be retroactive inhibition. They further mention Heine's technique as a possible explanation of her failure to get retroactive inhibition with recognition, pointing out that she held her *S*s to the criterion of regarding a syllable as recognized only if it gave a *distinct impression of familiarity*. The *S*s gave percentages exceedingly low for recognition experiments, probably meaning that they set a very high standard of recognition and gave "yes" answers only to impressions very firmly fixed and more or less proof against inhibition of the retroactive type. Whether this explains Heine's results or not, Woodworth and Poffenberger feel that we should suspend judgment on the Müller and Heine conclusion that recognition, since it is not susceptible to retroactive inhibition, is based on something other than association.

Spight (**24**), using 51 *S*s, compared two 12-hour rest periods between the partial learning and complete learning of lists of 14 paired words. One rest period occurred during the day and the other during the night, the major portion of the latter period being devoted to sleep. Each *S* learned 4 lists for each type of rest interval, each list apparently involving a separate pair of experimental sittings. Comparing the average number of repetitions required for complete

learning under both conditions, there was, between group averages, a difference of half a repetition in favor of the night rest (P.E. of diff. is 0.15), meaning that the learning with night rest took 9 per cent less time. Her work is open to the criticism of concluding from group averages only. No determination of the efficacy of sleep for each *S* is given. It is obvious that if sleep were detrimental for some *S*s and beneficial for others, the data would partially counterbalance, giving the results which she obtained. Other variables were pretty well controlled, except that there was only a partial control of that important factor, diurnal variation in learning performance. She concludes that a rest period, the majority of which is composed of sleep, is more effective for learning than a similar period of day activity. She points out that these results are in harmony with the findings of Jenkins and Dallenbach regarding sleep and retention.

Boreas (2), working at the University of Athens, has reported a study of the curve of forgetting, using adults of both sexes, 20 of whom were tested with lists of 15 nonsense syllables and 20 others with poetry, using the savings method. Every *S* learned only one list for every time interval, which of course does not give reliable results for any one given *S*. With both nonsense and meaningful material he found better retention after 24 hours than after 9 hours. With nonsense syllables the percentage of retention after 24 hours is 1.62 per cent more than after 9 hours. (See Table I.) With meaningful material 24 hours gave 8.69 per cent more retention than 9 hours; 48 hours gave slightly better retention than 24 hours (0.89 per cent); while 72 hours, though below 24 hours, gave 4.54 per cent more retention than 9 hours; 96 hours was practically the same as 9 hours; while the rest of the periods gave a gradually descending curve. Boreas interprets the greater amount of forgetting during the first 9 hours as due to fatigue resulting from the memorizing and the work of the day; although apparently he did not take account of diurnal variation in learning performance, which is a more plausible explanation of his result. The recovery at 24 hours for both types of material and the peculiar results in the case of poetry, are attributed to two factors: (1) rumination that has meanwhile occurred; and (2) to the deeper ingraving of the learned material in the brain resulting from physiological causes. It is just possible that the above mentioned failure to take care of diurnal variation may account for a good deal, if not all of the recovery reported on days following the learning. That is, the percentage of retention for the 9-hour period may be too low. This is, however, another study with

discrepancies similar to those of Ebbinghaus, Radosavljevich, and Luh, and possibly related to the question of sleep and retention.

van Ormer (27) has conducted a study to obtain further information regarding the effect of sleep upon retention. It was proposed in particular to repeat Jenkins and Dallenbach's work, comparing the retention of nonsense syllables after 1, 2, 4, and 8 hours of waking and of sleep. Employing the more precise and exact method of savings, it was hoped to obtain results, through refinement of technique, that would explain discrepancies between earlier investigations. Special care was employed in constructing 12-syllable lists; three lists were learned at one sitting; and 8 experimental periods for each time interval of both sleep and waking were conducted for each *S*, involving the learning of 24 lists for each point determined on the curves. Tests of diurnal and nocturnal variation in learning performance were conducted and corrections made in the retention scores were necessitated. Practice effect was carefully controlled. Two thoroughly practiced *S*s subordinated all activity to the conduction of this experiment.

The results are presented in Fig. II. Retention is definitely better after 8 hours of sleep than after the same time interval of waking; there is a probably reliable difference in favor of sleep at the 4-hour interval; a merely suggestive difference in favor of sleep at the 2-hour interval; and no difference between sleep and waking at the 1-hour interval. From the 1-hour period on, the retention curves for sleep are practically horizontal lines. The similarity in the curves of the 2 *S*s is significant. At the 1- and 2-hour intervals the three studies, Jenkins and Dallenbach's, Dahl's, and van Ormer's stand in opposition to one another. Remembering the drowsiness difficulty reported by Dahl and the report of Jenkins and Dallenbach's *S*s that at times they had no recollection of having been up the night before to do the recalling (*supra*, p. 13), van Ormer took especial pains to get his *S*s thoroughly awake before doing the relearning at night. It may be that the Jenkins and Dallenbach *S*s did some of their recalling in a trance state and accordingly did better than they would if they had been fully awake (see note 12). It should be noted that the *S*s tested in the two earlier investigations did not know when they were to be wakened for testing in the sleep experiments; while van Ormer's *S*s were told, at the conclusion of every initial learning period, the time when they were to be wakened. So far as the latter's *S*s could tell subjectively, the fact that they knew when they were to be wakened did not affect their going to sleep and therefore does not seem to be acceptable as even a partial explanation for the dis-

crepancy at 1 hour. In view of the great divergence in procedure, and in difficulty of method, the fact that the three studies agree in

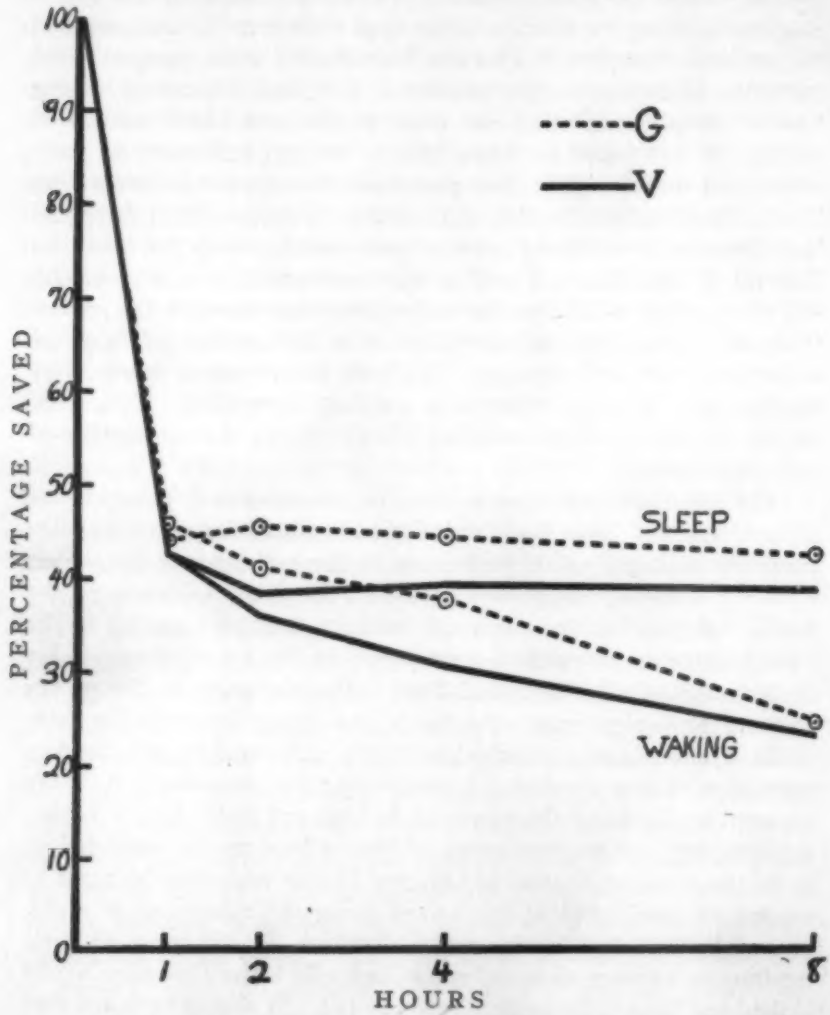

FIG. II. Average percentage saved for each *S* after the various time-intervals of sleep and of waking. (From van Ormer, *Arch. Psychol.*, 1932, 21, No. 137.)

confirming the conclusion that retention of nonsense syllables after 4 and 8 hours of sleep is definitely better than after 4 and 8 hours of waking, would seem to be of major importance.

van Ormer offers the following theoretical interpretation of his

results: "It is suggested that our results in favor of sleep are brought about by the absence of the inhibition and obliteration of the learned material by the waking activity. But if it is just a question of absence of inhibition, why do the curves of sleep and waking coincide at 1 hour and diverge from then on? May it not be that the waking activity not only inhibits and obliterates what has been learned, but that it also prevents or holds in check a perseveration or consolidation process which continues for a while in the nervous system after the impression of the learned material?[15] The perseveration process may frequently be at its height for the first part of the hour following learning.[16] The inhibition produced by the waking activity for the first ten minutes after learning, before it is possible to get to sleep, and possible inhibition from the process of going to sleep, itself, thus interfere with a very effective part of the perseveration process; and accordingly their effects practically overbalance, for the 1-hour sleep period, the 50 minutes of sleep within which there is a possible continuation of the perseveration process and the absence of inhibition and obliteration by waking activity. It is also very possible that the process of getting awake and the little activity interpolated between awaking and relearning also have an inhibitory effect which would aid in keeping the retention for the 1-hour period down to the level of the waking period. These results on the 1-hour period would suggest that a perseveration process does exist, at least for the first 10 or 15 minutes following learning.

"The results in favor of sleep for the remaining time intervals, and the results of the Jenkins and Dallenbach study from 2 hours on, are hard to reconcile with the law of atrophy through disuse. It may be argued that sleep refreshes the organism and recall is thereby benefited. But if such is the case it may be asked why retention, or reinstatement, after 8 hours of sleep is no better than after 1 hour of sleep? Certainly refreshment or physiological recovery—if produced by sleep—is far from complete after 1 hour. And it must be remembered that the retention after 1 hour is the same for sleep and waking in the case of our *S*s, which would hardly be the case if the sleep had produced a refreshed state which was beneficial to reinstatement. The more plausible explanation of our results in favor of

[15] This idea of a perseveration process is not new. See, *e.g.*, Müller and Pilzecker (18) and Skaggs (23), who have argued for a perseveration process to explain retroactive inhibition. Woodworth (29) has also argued for a consolidation process following active learning.

[16] Skaggs (23, p. 27), "Work introduced immediately after the learning produced the greatest amount of inhibition." Also see Skaggs (22).

sleep for the latter three periods would seem to lie in the further absence of inhibition and obliteration of the learned material by the waking activity; with the additional possibility that the perseveration process may continue for a longer period than the first hour.[17] . . . Our results, . . . would suggest that a primary factor in forgetting is the action of the interpolated activity as it inhibits a consolidation or perseveration process and produces inhibition and obliteration of the learned material" (27, pp. 45–46).

A striking side result of van Ormer's study is his failure to find any daytime variation in learning performance for one *S* and the failure of 5:30 P.M. to show any variation from 9:30 A.M. in the case of both *S*s. This result fails to confirm the result of Ebbinghaus, who found his performance at 6–8 P.M. to be 12 per cent of itself below that at 10–11 A.M.; and is against the argument put forth, as we have indicated above, by Radosavljevich, Bean, Finkenbinder, and Boreas, that fatigue would cause relearnings at 5:30 P.M. to take longer and consequently cause a drop in the retention curves at the 8-hour interval if those relearnings took place at 5:30 P.M. without corrections for diurnal variation. There are probably definite individual differences in diurnal variation of ability to learn nonsense syllables.

Hunter (12) has recently reported an investigation, suggested by the Jenkins and Dallenbach experiment, in which he studied the effect of inactivity produced by cold upon the learning and retention of a simple darkness avoiding response by the cockroach. The roaches were conditioned by an electric shock received when they entered the dark side of a box. Thirty to forty different *S*s were used in each section of the experiment. All conditions were carefully controlled.[18]

[17] Note in this connection the tendency of interrupted activities to perseverate. See, *e.g.*, Zeigarnik (31). It is also relevant to note the tendency of the day's activities to perseverate some hours later, or as one lies in bed before going to sleep. Woodworth (29, p. 92) has suggested the possibility of a consolidating process lasting for some *hours* following active learning.

[18] Hunter's article does not seem to indicate whether diurnal variation of learning performance was controlled. In a letter to the writer under date of March 17, 1933, Professor Hunter says, "The experiments began at 9 A.M. and extended without intermission until 9 P.M. All groups had an equal number of animals tested, both for learning and relearning, at every hour of the day—except that Group c could not begin until the requisite number of hours after 9 A.M. Groups a, b, and c were always carried along simultaneously. First an animal from *a*, then one from *b*, then one from *c*—then abc, abc, etc., until the requisite total number in each group was secured. Group I gives an additional check on diurnal variations, and none was found."

He concludes from statistically reliable differences that the learning of the darkness avoiding response is retarded if the roach is exposed to low temperatures (3°–6° C.) for 2 or 4 hours followed by a period of recovery of 1 hour in normal temperatures prior to the learning. Using the relearning method, retention of this response for roaches living in normal temperatures was found to decrease from 1 hour to 8 hours elapsed time, there being a nearly perfect retention during the first hour. If 2 hours of interpolated cold and 1 hour of recovery at normal temperatures were employed, neither better nor worse retention was obtained than under normal conditions. Four hours interpolated cold and 1 hour recovery under normal temperatures; and 4 hours cold with 4 hours recovery were unfavorable to retention, but in equal amounts.

Hunter had thought that almost perfect retention might be secured if the roaches' physiological processes could be held practically in quiescence during the period of rest interpolated between learning and relearning. He points out that the experiment's failure to fulfill the expectation does not negate the expectation in relation to other forms of quiescence. It is not necessary to expect all forms of inactivity to have the same effect, or lack of effect upon retention, since physiological accompaniments of inactivity may well vary with the mode of producing the inactivity. Inactivity due to cold may not have the same relation to retention as inactivity due to sleep. Thus he accounts for the discrepancy between his results and the former studies of sleep and retention. He points out that the experimental data offer no explanation of the facts; but that histological studies might have revealed that definite structural changes had resulted from the low temperature and persisted at least 4 hours after shift of temperature to normal again. He cannot say what physiological changes were produced by the temperature used, other than a general slowing down of organic activity.

Conclusions

(1) In consideration of the studies of Heine, Jenkins and Dallenbach, Dahl, and van Ormer, the conclusion seems established that the usual amount of daily sleep (about 8 hours) favors retention of nonsense syllables over that time interval.

(2) Regarding the more specialized generalization that retention of nonsense syllables is better *even after short time intervals of sleep* than it is after the same intervals of waking, the results are conflicting. Jenkins and Dallenbach's, Dahl's, and van Ormer's results warrant the conclusion that retention of nonsense syllables is better

after 4 hours of sleep than it is after 4 hours of waking. The first and last of the three studies suggests that the advantage of sleep is not as great at 4 as at 8 hours. It is apparent that more study of the 1- and 2-hour periods is desirable; especially should Dahl's study by the recognition method be repeated and every effort made to have the *S*s thoroughly awake when they are tested.

(3) Jenkins and Dallenbach suggest, and van Ormer concurs with them, that the favoring of retention by sleep accounts for Ebbinghaus' discrepancy from 8.8 to 24 hours. Sleep would also seem to be a partial explanation for the discrepancies in the curves of Radosavljevich, Boreas, and the relearning curve of Luh. These investigators, not controlling diurnal variation in performance, probably have it as at least one factor to explain their discrepancies.

(4) The problem has further ramifications. If retention of nonsense syllables is better after 8 hours of sleep than after 8 hours of waking, how long, it may be asked, does this greater retention persist? Or, in other words, what happens to the shape of the retention curve from 8 hours on? Heine's study offers some data on this question, since she found that retention of nonsense syllables after 24 hours was better when the learning had occurred just before retiring, than when intervals of waking activity were interpolated between the learning and sleep. Further investigations on this phase of the problem are important.

(5) In view of some fundamental similarity between the learning of nonsense syllables and other verbal material,[19] the suggestion seems to be warranted that conclusions (1) and (2) may apply to other verbal material. Accordingly, the four studies mentioned in (1), combined with Spight's suggestive results, indicate that retention of any verbal material is probably better after 4 or 8 hours of sleep than after the same time intervals of waking. Further study with meaningful material is very desirable.[20]

(6) If conclusion (5) is thoroughly established, a new principle of efficient study will be clearly indicated. Assuming that a possible decrease in the learning *efficiency* at the late evening hours is overbalanced by the better retention following sleep, the advantage of

[19] Woodworth and Poffenberger (30, Chap. 6, Sect. VII) analyze various memory experiments, showing that nonsense syllable learning is not a mere mechanical process, but an observant, active looking for groupings and relationships, such a process as one carries on in the learning of other verbal material.

[20] Studies with motor performances are also of practical and theoretical value. van Ormer (27, p. 44) has listed a number of problems for further study of sleep and retention.

night study, as suggested by Jenkins and Dallenbach, will be evident. Much less is forgotten during sleep, and "on waking the learner may take up the task refreshed and with renewed vigor."

(7) The results of the studies on the effect of sleep upon retention are of great theoretical importance. They have an important bearing upon the much discussed subject of retroactive inhibition; and combined with recent work on this latter subject, indicate, as McGeoch (17) has pointed out, that the accepted theories of retention and of forgetting will have to be considerably modified. The studies on sleep and retention provide a crucial experiment in criticism of the "law" of atrophy accompanying disuse of the modifications produced in the nervous system by learning. The studies of sleep and retention and of retroactive inhibition seem to demonstrate that forgetting is to a large degree a function of the interpolated experience. Foucault (8) (*supra,* p. 4) and Jenkins and Dallenbach (13) (*supra,* p. 11) were led to postulate this theory independently. Robinson (21) has accepted it, and explains the loss of habits and ideational material as due to altered stimulating conditions and to the interference of interpolated activity. Hunter (10) in 1928, accepting this theory to some extent, explained forgetting as follows: "A given form of behavior is absent (forgotten) at the present moment either because of the absence of the proper stimulus, because of the interference of other habits or neural processes, or by virtue of disuse which has rendered reinstatement difficult" (p. 293). In 1929, Hunter (11), in reviewing the work on retroactive inhibition and the work of Jenkins and Dallenbach, agrees with their conclusion regarding forgetting (*supra,* p. 11) and suggests, as did Foucault twenty years ago (*supra,* p. 5), that it is probable that such experiments, involving interpolated intervals of sleep, would show no rapid initial drop if it were possible for the *S* to go to sleep *immediately* after learning. McGeoch (16, 17), in criticizing the law of disuse and in interpreting various facts related to forgetting, has set forth the hypothesis ". . . that forgetting is a function of the kinds and conditions of interpolated experience and of the nature and degree of the shift in stimulation, not of the period of disuse as such. Other factors . . . may set limits to the operation of these primary factors, but are themselves secondary in importance" (16, p. 670). "Disuse is important only in that it so often gives the primary conditions an opportunity to act. With further research it may be possible to bring the two major conditions mentioned under a single concept" (17, p. 370). van Ormer has suggested, ". . . it is quite possible that forgetting is a function of the nature of the

interpolated experience, the altered environmental contexts, and the organic state of the individual during and subsequent to learning, rather than intrinsically a function of the period of disuse" (27, p. 47).

As Carr (3) has pointed out in discussing the nature of forgetting, we must not lose sight of the two-fold explanation of the phenomenon of inability to reinstate an activity. He argues that forgetting is sometimes to be explained in terms of defective retention because the organic effects that condition the reinstatement of an activity have undergone a process of disintegration due to the ravages of time, onset of senility, nervous shock, use of drugs, and effects of disease; and that in other instances forgetting may be due to inability to recall rather than to any defect of retention, the factors here being, first, lack of associations with the present situation, either because such associations were never established or have become dissociated, and second, the repression of a given material by later acquisitions.

The writer would like to suggest the following brief tentative interpretation of forgetting. When it is purely a matter of inability to recall, temporarily or for long periods of time, even though the structural modification, or " memory trace," still remains in the organism, the effective factors would seem to be altered environmental contexts, interference produced by later acquisitions (transfer theory of retroactive inhibition), inhibition of the acquired activity through the establishment of a stronger system of associations attached to the same situation, and the organic state (emotional state, disease, hypnosis, influence of drugs) at the time of attempted reinstatement. In spite of cases of unusual recall of childhood acquisitions,[21] there is no evidence for the extreme view that anything once learned is never forgotten.[22] Accordingly it seems plausible to assume a loss of retention which is the result of a partial or complete disintegration of the neural trace or special structural modification produced by the learning. This might first of all be due, basically, to a physiological atrophy accompanying disuse of the modifications produced in the nervous system by learning. The writer has suggested elsewhere (27) (*supra*, p. 19) that inhibition of a consolidation or perseveration process immediately following learning is an important factor in forgettng. Inhibition of such a process would prevent the thorough fixation of a structural modification and might

[21] *Cf.*, for example, Carr (3, p. 246).

[22] Woodworth (29, p. 93–94) has pointed out the folly of such a view.

make it more susceptible to physiological atrophy. It is also conceivable that interpolated activity could partially obliterate the structural modification produced by antecedent activity, thus making the partially disorganized structural modification more susceptible to physiological atrophy. As a second basic factor, it is also possible that at times interpolated activity may completely obliterate, or at least practically obliterate, so far as the possibility of any functional activity is concerned, the structural modification produced by antecedent learning. Another factor which may operate by speeding up, or slowing down, the action of physiological atrophy, is the organic state (emotional state, hypnosis, sleep, drugs, disease, senility) during and subsequent to learning. Hunter, in discussing the experiments on sleep and retention, says, "The evidence, however, does not enable us to say that the neural impulses set up by muscular and glandular activities are the only important factors at work (in producing forgetting), since during sleep the metabolic processes of the brain may well be different from their condition during waking" (12, pp. 253–254). It is conceivable, though in the writer's opinion not so probable, that during sleep the physiological atrophy may be balanced by a counter physiological building up process, gradually increasing in effect as the atrophy does, and thus producing the horizontal curves of the Jenkins and Dallenbach and van Ormer experiments. As indicated above (*supra*, p. 19), such a theory would seem to call for a rise in the retention curve after longer intervals of sleep instead of the horizontal line obtained in the two investigations.[28] In conclusion, suffice it to say that the whole problem is a complex one. The statement that the rate of forgetting diminishes with time is hardly correct as an unqualified statement; and the explanation that forgetting is produced by the physiological law of atrophy through disuse is far from complete.

BIBLIOGRAPHY

1. Bean, C. H., The Curve of Forgetting. *Arch. Psychol.*, 1912, No. 21.
2. Boreas, Th., ΠΕΙΡΑΜΑΤΙΚΑΙ ΜΝΗΜΟΝΙΚΑΙ ΕΡΕΥΝΑΙ, Β. Η ΠΟΡΕΙΑ ΤΗΣ ΛΗΘΗΣ *Πρακτικῶν τῆς Ἀκαδημίας Ἀθηνῶν*, 5, 1930, σ 382. (Experimental Studies in Memory. 2. The Rate of Forgetting. *Proceedings of the Academy of Athens*, 1930, 5, 382.)
3. Carr, Harvey, *Psychology*. 1925. Pp. 432.

[28] The writer is indebted to Dr. M. N. Chappell, who has suggested this physiological conception through conversation and an as yet unpublished manuscript, "The Rôle of Activity in Retention." Chappell (4) has expressed a physiological conception of the typical retention curves.

4. CHAPPELL, M. N., Chance and the Curve of Forgetting. *Psychol. Rev.*, 1931, **38**, 60–64.
5. DAHL, A., Über den Einfluss des Schlafens auf das Wiedererkennen. *Psychol. Forsch.*, 1928, **11**, 290–301.
6. EBBINGHAUS, H., *Ueber das Gedachtnis.* 1885. English translation by H. A. Ruger and C. E. Bussenius, Memory. 1913.
7. FINKENBINDER, E. O., The Curve of Forgetting. *Amer. J. Psychol.*, 1913, **24**, 8–32.
8. FOUCAULT, M. MARCEL. Introduction à la psychologie de la perception: Expériences sur l'oubli ou sur l'inhibition régressive. *Rev. des Cours et Conferences*, 1913, **21**, Series 1, No. 5, 444–454.
9. HEINE, ROSA, Über Wiedererkennen und rückwirkende Hemmung. *Zsch. Psychol.*, 1914, **68**, 161–236, esp., 225–234.
10. HUNTER, W. S., *Human Behavior.* 1928.
11. HUNTER, W. S., Chapter 15, in *Foundations of Experimental Psychology.* 1929. C. Murchison, ed.
12. HUNTER, W. S., The Effect of Inactivity Produced by Cold upon Learning and Retention in the Cockroach, Blatella Germanica. *Ped. Sem. & J. Genet. Psychol.*, 1932, **41**, 253–266.
13. JENKINS, J. G., and DALLENBACH, K. M., Obliviscence During Sleep and Waking. *Amer. J. Psychol.*, 1924, **35**, 605–612.
14. LUH, C. W., The Conditions of Retention. *Psychol. Monog.*, 1922, **31**, No. 3, Whole No. 142, 1–87.
15. MCDOUGALL, W., Experimentelle Beiträge zur Lehre vom Gedächtnis. vom G. E. Müller und A. Pilzecker, *Mind*, 1901, **10**, 388–394, esp. 393.
16. MCGEOCH, J. A., Forgetting and the Law of Disuse. *Psychol. Bull.*, 1931, **28**, 670. (Abstract of paper read to 1931 Amer. Psychol. Assoc.)
17. MCGEOCH, J. A., Forgetting and the Law of Disuse. *Psychol. Rev.*, 1932, **39**, 352–370.
18. MÜLLER, G. E., and PILZECKER, A., Experimentelle Beiträge zur Lehre vom Gedächtnis. *Zsch. Psychol.*, 1900, Erg. 1, pp. 300.
19. NICOLAI, F., Experimentelle Untersuchung über das Haften von Gesichteindrücken und dessen zeitlichen Verlauf. *Arch. ges. Psychol.*, 1922, **42**, 132–149, esp. 137.
20. RADOSSAWLJEWITSCH, P. R. (now Radosavljevich), Das Behalten und Vergessen bei Kindern und Erwachsenen nach experimentellen Untersuchungen. *Päd. Monog.*, 1907, I.
21. ROBINSON, E. S., *Practical Psychology.* 1926. Pp. 479, esp. 129, 281–282.
22. SKAGGS, E. B., Further Studies in Retroactive Inhibition. *Psychol. Monog.*, 1925, **34**, No. 161, 1–69.
23. SKAGGS, E. B., *The Major Forms of Inhibition in Man.* 1931. Pp. 104.
24. SPIGHT, JULIA, Day and Night Intervals and the Distribution of Practice. *J. Exp. Psychol.*, 1928, **11**, 397–398.
25. STALNAKER, J. M., and RIDDLE, L. E., The Effect of Hypnosis on Long-Delayed Recall. *J. Gener. Psychol.*, 1932, **6**, 429–440.
26. STRONG, E. K., JR., Two Factors which Influence Economical Learning. *J. Philos., Psychol., and Scient. Method.*, 1914, **11**, 124–131, esp. 130–131.

27. VAN ORMER, E. B., Retention After Intervals of Sleep and of Waking. *Arch. Psychol.*, 1932, **21**, No. 137.

28. VAN ORMER, E. B., and DALLENBACH, K. M., A Frequent Error Concerning Ebbinghaus' Experiments on Obliviscence. *Amer. J. Psychol.*, 1931, **43**, 706–707.

29. WOODWORTH, R. S., *Psychology*. Revised 1929. Pp. 590.

30. WOODWORTH, R. S., and POFFENBERGER, A. T., *Textbook of Experimental Psychology*. 1920. Mimeographed edition in Columbia University Library. Pp. 285.

31. ZEIGARNIK, B., Das Behalten erledigter und unerledigter Handlungen. *Psychol. Forsch.*, 1927, **9**, 1–85.

BOOK REVIEWS

TOLMAN, EDWARD CHACE. *Purposive Behavior in Animals and Men.* New York and London: The Century Psychology Series, The Century Co., 1932. Pp. xiv+463.

Tolman's book is not only a most stimulating piece of work, it may also become a milestone and signpost on the road of American psychology. But it is a difficult book to review, particularly if the reviewer is of the opinion that the achievement of the author is not to have taken his readers to a goal—nobody would expect this, and the author would be the last to raise such a claim—but not even to the end of a stage. The reviewer sees its importance in the fact that the author has turned a corner, he cannot help hoping that the author's coach is still on the road, and that this road, if steadily pursued, will lead him to a destination at which at the present time he seems loth to aim.

Tolman calls himself a purposive behaviorist, he sees the object-matter of psychology in what he happily terms *molar* behavior. This definition of the task of the book explains his method and his theoretical equipment. It contains all the difficulties which the author tries to conquer and with which his readers are faced. For it means the combination of two problems: the proof that behavior is molar, and the claim that psychology can and must be behavioristic, provided that behavior is defined as molar behavior. The reviewer need not say that he agrees without reservation to the first part, but he will add that in his opinion Tolman has produced a great mass of evidence, taken from beautifully planned experiments, for his contention. But the reviewer feels that the second claim has prevented the author from making the most of his results by encumbering them with a system of concepts, which despite many efforts the reviewer has not been able completely to understand.

What does Tolman mean by molar behavior? He opposes it to molecular behavior, *i.e.*, behavior "in terms of simple stimulus-response connections" (p. 4), "in terms of its strict underlying physical and physiological details, *i.e.*, in terms of receptor-processes, conductor-processes, and effector-processes *per se*"(7). Against this definition he maintains that "behaviour, as such, is more than and different from the sum of its physiological parts. Behaviour, as

such, is an 'emergent' phenomenon that has descriptive and defining properties of its own. And we shall define this latter as the *molar* definition of behaviour" (*7*). Behaviors *qua* molar are, for instance, a rat running a maze, a man driving home to dinner, a woman gossiping over the telephone (8). Molar behavior has three descriptive properties, it is directed towards or away from a specific goal, it requires commerce with intervening means objects, and it selects the best, shortest, easiest, means object which will lead it to its goal, or otherwise expressed, it is both purposive and cognitive.

Two questions arise at this point: what is the relation between molar and molecular behavior, and what place can purpose and cognition have in an objective, behavioristic psychology? Our author tries to answer both, but right here the difficulties begin.

In answer to the first question he states "that 'behaviour acts,' though no doubt in complete one-to-one correspondence with the underlying molecular facts of physics and physiology, have as 'molar' wholes certain emergent properties of their own" (*7*). And soon after he adds that "behaviour as such cannot, at any rate at present, be deduced from a mere enumeration of the muscle twitches, the mere motions *qua* motions which make it up" (8). It must then be the opinion of the author that molar behavior is in complete correspondence with physiological processes which *in themselves are molecular,* and that possibly in a distant future our knowledge of these molecular facts will be sufficient to explain molar behavior, even if at the present time we have to turn to molar behavior directly. If this is the correct interpretation—and I cannot think of any other compatible with the author's own words, and find it confirmed by his theory of the conditioned reflex, which he rejects as a molar, but accepts as a possible molecular phenomenon (336)—the relation between molar and molecular behavior remains a mystery to me. For the molecular facts are primary realities; molar facts are "emergent," and yet in one to one relation with molecular facts, which in the long run may replace them completely. What is, in this theory the status of reality of molar behavior? I have asked myself this question again and again in reading Tolman's book, and I must confess that I have found no answer to it.

The answer to the second question helps little to clarify our issue. The purposes and cognitions immanent in all molar behavior are defined as functional variables. Purpose is defined by persistence through trial and error, and docility; and cognition is defined by docility with regard to the character of the goal object and its initial

position, and the characters of the intervening means-end objects. Thus purpose and cognition are "objective and it is we, the outside observers, who discover—or, if you will, infer or invent—them as immanent in, and determining behaviour. They are the last and most immediate causes of behaviour. We call them therefore 'immanent determinants'" (19). I have great difficulty with this theory. These determinants are called objective, and that means, I believe, they are not "mentalistic," not concepts referring to the experience of the animal. On the other hand Tolman does not object to their being called "inventions of the outside observer's." If they are that, are they not then subjective? And if our author has succeeded in divorcing them from the experience of the observed animal, has he not at the same time incorporated them in the experience of the observing human?

I believe that many readers will feel as I do that this whole difficulty is created by Tolman's wish to be a behaviorist as well as a purposive one. He is so much afraid of falling into the snares of mentalistic theories that he sacrifices the concreteness of his own theory without being able to escape "mentalism," and at the risk of inconsistencies. For, if these determinants are mere inventions of the human observer, they cannot be causes of molar behavior, unless molar behavior is itself only an invention. Sometimes it seemed to me as though Tolman's system dwelt in a realm of pure fictions, of *as-ifs,* connected with reality only by their power of suggesting experiments. And since in the last words of his book he confesses himself a radical pragmatist, this appearance may not be entirely false. But it is not the whole truth either; otherwise he would not defend purposive molar behavior with so much pluck and enthusiasm. Also does he not tell us merely that we invent these determinants, he gives us two more words to choose from, discover and infer. If we accept these, which to my mind is equivalent with rejecting the first, we can understand how these discovered or inferred factors cause observed consequences in a real world. Then they are truly objective, and their locus is the animal which we observe. But this brings us back to our first difficulty, the relation between molar and molecular behavior. If a one to one correspondence exists between them, then every piece of molar behavior has a counterpart in molecular behavior. Molecular behavior is fully explained by molecular causes and these, so it seems, also account for molar behavior in which nothing actually new happens in the sense that it had nothing corresponding to it in molecular behavior. Again, what we believed we had gained has slipped from our grip.

And the same will o' the wisp nature of his molar concepts manifests itself, in my opinion, in their definition. Throughout, Tolman identifies explicitly criterion and definition, a practice, I believe, widely used in modern logic. One quotation taken from a great number of similar ones: "The point we would make is this that such 'searching,' *i.e.*, disruption, is to be taken as the empirical evidence for, and definition of, an immanent expectation of the previously obtained bran mash" (74). Now I admire Tolman's skill in devising criteria which can be used in experiments, but I do not agree with him that these criteria can at the same time serve as definitions, provided we are not dealing with mere fictions. To accept a certain phenomenon like searching or docility as a criterion for the existence of a certain trait of molar behavior means to consider it as a sufficient condition for our inference. Whenever the particular criterion is fulfilled we are justified in assuming the real existence of the particular trait. But this is not a definition in the sense of describing the actually existing trait. The excursion of the pointer on a seismograph is the sufficient condition for my inference that an earthquake has occurred; but it does not define an earthquake. By a mere fiat Tolman makes his sufficient condition also a necessary one. This would be justified in a game the rules of which he could lay down. If we accept it, therefore, we must consider his whole system as a structure of fictions; we should only be surprised to find names well known to us, like purpose, expectation, cognition, applied to such arbitrary as-ifs, just as we find kings and queens and knights in a game of chess. But then we know that these are not real kings or queens. I do not believe that our author merely wants to play such a game. After all, why did he establish his criteria? I suppose he started from his previous knowledge, which he shares with every one of us, of what, *e.g.*, an expectation is. Then he attempts to translate such knowledge into a behavioristic system. There is nothing reprehensible in this. On the contrary, it is very important to find observable behavior traits characteristic of expectation, and thereby to establish reliable criteria for the existence of expectation. The error begins, in the reviewer's opinion, when the purist enters and says: Now, after having discovered such criteria, let us discard the object of which they are criteria altogether, for only the criteria but not the objects are observable. But this leads to absurdities. Return to our last quotation; expectation is there defined by a searching behavior, or more objectively, by a disruption of behavior in the food box. On the other hand, this expectation is one of the causes

which send the rat through the maze, and as such a cause it is operative at the moment the rat leaves the starting box. It is of course nonsense to say that the disruptive behavior in the food box causes the behavior in the starting box, and yet that is what we should have to do, if the criterion of expectation were taken for its definition. Expectation has many effects, some of which, like the disruptive behavior, can be used as criteria for its existence, but expectation is not the same as any of its effects. Just as little as we define an earthquake by the operation of the seismograph can we define expectation, purpose, or any of the other concepts of molar behavior, by criteria which are but effects of particular aspects of behavior. Tolman, I am certain, wants to establish the reality of these molar characteristics. But being a pragmatic behaviorist, he substitutes for these realities mere fictitious constituents of a logical system.

In two points, it seems to me, he will have to modify his basic structure in order to clarify his system. In the first place he must abandon all modesty and claim that behavior is molar all through, and that molecular facts exist only either in the artificial nerve preparation, or for the physiologist who studies a small part of the system and ignores for the moment its connection with the rest. In reality the physiological processes within the animal are as molar as his molar behavior. With this assumption, which by now is a good deal more than a mere assumption, a number of difficulties disappear at one stroke. Then all molar concepts refer to actual events within the animal.

But to make this solution really satisfactory, the author would have to take one more step; he would have to revise his behaviorism. Purpose, cognition, expectation, are characteristics of behavior. A rat runs for bran mash and not for sunflower seed, it runs in a certain direction and not in another, it has a general orientation in the maze. What can such descriptions mean? The goal is not the goal as the experimenter knows it who has put it in the food box; to him it is not a lure. Just so are the cognitions and expectations the rat's cognitions and expectations; the alley walls through which it moves are *its* alley walls, high and forbidding, and not the experimenter's alley walls, small and easily handled. In other words, what Tolman needs is a frank recognition of the animal's environment as something made by or within the animal, and which determines its behavior. In such a "behavioural" environment all his concepts fall into their proper places; this environment is the field of which he speaks himself (176 f), and this field can be explored, but not defined, with the help of his criteria.

How, then, is it to be defined? Let us for the moment abandon the rat as our subject and substitute Mr. Tolman while he makes an experiment with a rat. What is true of the rat's behavior is equally true of his; it is behavior in a field, and this field consists of the maze he has set up, the rat he has put into it, the activities of the rat, the previous knowledge he has about his problem and the interest he has in it. This field is realized only once, viz., at the time Mr. Tolman makes this experiment. Completely it can never be repeated, as no moment of the universe is ever quite repeated. It is even private, for nobody can know this field as well as Mr. Tolman, nobody will see the experiment quite in the way he sees it. And yet, all his conclusions are due to this private field of his. Since this field is the foundation of his whole knowledge, since only by behaving in a great number of such fields could he have written his book, he cannot pretend that this field does not exist. No mentalist needs to assert more than the existence of such fields, no more need he mean by "consciousness." And if no more is meant by this term, widely misused, and for that reason rightly disliked by the author, then psychology, as the study of behavior cannot do without consciousness. For at least the field of man is a conscious field. Strictly speaking, of course, I can only call my own field a conscious one, it is the only one I know. But nobody is a solipsist.

Let us return to the rat. Can we ascribe similar conscious fields to this alert rodent? Tolman might answer that we have no criterion whatever to decide whether we can or not. But is the same not true for our fellow men? Exactly, Mr. Tolman will reply, and that is my reason for discarding this whole concept. Only, I must remind him, you cannot discard your own field. When you speak of rats and mazes, of searching or exploratory behavior, you are all the time referring to your own conscious field. So either you must claim that you are unique in the world, or you must be more generous and admit that what belongs to you may also belong to other beings and play the same rôle in their behavior which it plays in yours. I don't know what Mr. Tolman's answer to this argument will be. In case he accepts it, however, he could still come back by asking: what good does it do to psychology to assume consciousness in rats when we have no reason whatever for this assumption except an idea that we are not unique creatures. To this question I would answer: let us discard the conscious field of the rat; we still must keep its field, for in its field it behaves. This field must then be a molar physiological field. But why then my insistence on our conscious fields? Simply, because directly we know nothing of our own brain fields, but a great

deal about our conscious fields, and because we theoretically construct the characteristics of our brain fields in such a way that they fit our observed conscious fields. The brain field must, *e.g.,* be such that it corresponds to seen objects with their particular "demand characters," in their particular spatial and temporal distribution, and so forth. In other words, we must describe the brain fields with reference to and in terms of our observed conscious fields. We can do nothing else when we speak of rats. In describing their brain fields we cannot help using terms which refer primarily to conscious fields, even if we abstain from ascribing conscious fields to rats. Still their brain fields must be such that they would correspond to certain conscious fields and not to others, provided they had conscious fields. Tolman's own terminology is decidedly of that nature. He tries to escape our conclusion by his identification of criterion and definition; but, as we have shown before, that method does not work. Let us then take purpose, cognition, expectation, at their face values, as actual processes within the rat, which, if they were accompanied by consciousness, would carry to the animal purpose, orientation, alley walls, generically, though not specifically, as we know them. All of Tolman's experiments retain their value, telling us a great deal about the specific nature of the rat's fields, about those which distinguish them from our own as well as about those in which they are similar to ours.

Thus I find we should escape all the difficulties which Tolman sees in the concept of consciousness, to which he devotes a whole chapter (XIII), which seems to me the most unsatisfactory of the whole book. In this chapter he introduces consciousness behavioristically as something which occurs at certain moments in the rat's behavior. But since he warns us against taking this speculation of his seriously (204), I shall not go into it any further.

Let us now go into some of the details of his system, his theory of behavior and of learning. Behavior, being purposive, involves expectations, a claim amply substantiated by fine experiments. Expectations "are always forward pointing" (84), they are fulfilled or disappointed by commerce with means-end- and goal-objects. "This statement of the duality of stimulation and verification is but another way of saying that any behaviour-act requires not only stimuli to release it, but also, later, more substantial environmental actualities such as *fulcra, media* and *planes* to *support* it (verify it, make it possible). Stimuli by themselves are not enough; *supports* also are needed. Behaviour cannot go off in vacuo" (85). There are two

kinds of such supports, the *discriminanda and manipulanda,* "broadly distinguished as 'sensory' supports on the one hand, and 'motor' supports on the other" (85). Their final definitions can be accomplished only by a whole series of experiments (86 ff.). The color pyramid is an example of an experimentally determined system of discriminanda, and similar representations should be found for the different manipulanda systems. Behavior supports are complex, inasmuch as they consist "of a sign object (*i.e.,* means object), a signified object (*i.e.,* goal object), and a signified means-end relation (direction distance correlation between such sign object and signified object)" (453 and 134 f.). This complex is called a *sign-gestalt,* of which there are three modes, perception, mnemonization, and inference (141). All learning consists in the building of sign-gestalt-expectations. Vast experimental evidence is adduced to support these theoretical elaborations.

This might be interpreted in harmony with the reviewer's previous criticism to mean: Behavior takes place in a behavioral environment—which is not the same as the real environment—which possesses a number of characteristics, demand-, manipulanda-, goal-characters. These characters determine the animal's behavior. In learning these characters change, and thereby the behavior of the animal. It is a characteristic of many—the author would probably say all—environmental objects that they are forward pointing, and likewise the behavior is forward pointing before the goal is reached, it is not only what it is at the moment, but also what it is to become at later moments. The whole process is a gestalt process, directly determined by the dynamic properties of the total field.

If this interpretation were correct, Tolman would have done what we said before he should have done. But is it? There are passages in the book which seem to confirm it. "Behaviour is . . . always ultimately an s→o→R phenomenon, or rather an s→(o)→R phenomenon, and not a simple S→R phenomenon." This is said in explanation of the support concept (336). This formula looks identical with Köhler's formula: "Constellation of stimuli—Organization—Reaction to results of organization" (Gestalt Psychology, p. 180). But, of course, in Köhler's theory this formula has a dynamic meaning; organization is not a fictitious concept, but means an actual interplay of real forces. Whether, therefore, Tolman's theory of supports can be interpreted in this way depends upon the interpretation which we must give to his general theory, and we have discussed the difficulties which beset this interpretation.

Tolman's own words as just quoted (from p. 85) do not, for me, dissipate the obscurity. When he says that behavior acts require *not only* stimuli to release them, *but also* environmental actualities to support them, he seems to treat stimuli and organizations produced by stimuli on an equal basis with regard to behavior. And again and again he uses the term stimulus, where Köhler would speak correctly of organization produced by stimuli. Moreover, he fails to distinguish between the *distant stimuli,* the stimulus objects, like the walls of the maze, or the food in the food box, and the *proximal stimuli, e.g.,* the retinal images, which the distant stimuli produce on the sense surfaces of the animal. These latter, and not the former, are in direct relation with the processes of organization. This oversight is very serious, for it prevents a proper formulation of the problem, what is the relation between stimuli and supports, or, in our terms, between stimulus constellation and organization. "Secondly, however, the organism must also have the capacity for being endowed with, or for acquiring, the correct expectational relationships . . . between the given sets of immediately presented stimuli and the to-be-expected discriminanda and manipulanda. . . . We want to know, for example, not only how well an animal can see under normal, standard, conditions—*i.e.,* the facts concerning his discriminanda capacities—but also how well he can *expect* the true color, size or what not, from given stimuli when the distorting factors of perspective, and the like, enter in" (91/2). In the first place a relation (expectation) between stimuli and discriminanda (the true color, etc.) seems to me to be a relation between objects which belong to different universes of discourse; in the second place the author seems to assume that stimulus constellations may be "not-distorted." But I can, *e.g.,* see a light grey color under stimulus conditions which vary over an enormously wide range, so that in three cases in which I see the same grey the stimulus intensities may be 1, 100, 1,000 in some arbitrary unit. Which of these three is the "standard condition"? From Tolman's point of view the solution of the constancy problems in perception is absolutely impossible.

Furthermore, even if we interpret his theory as a form of gestalt theory, it has a distinguishing feature, viz., what I could call its oversimplification. Tolman admits only one kind of gestalt, the sign-gestalt (421). Moreover, as the prototype of the sign-gestalt he seems to establish the case of the honk of a motor car which "will 'set' all but a new-born infant . . . for the whole signified complex: automobile plus impatient driver" (139). This, according to

gestalt theory, is a very weak gestalt, containing very little of the main gestalt characteristics. True enough, in his inferential sign-gestalt expectations (140, 164 f., 176) he allows for another and better type of sign-gestalt; but, from my point of view, at least the emphasis is wrong. Apart from that, the restriction of the gestalt concept to the sign function is quite unwarranted. It makes Tolman's theory a system in a sense in which gestalt theory is not. Gestalt theory raises the formal claim that the part-whole relationship of traditional psychology is but a limiting case of a great variety of relationships, which are all distinguished by the characteristic that the parts are parts in organization. But it does not say: all organization is of one particular type. And this is just what Tolman maintains, and therefore he has quite frequently to follow the dictates of his system rather than the evidence of the facts. A psychological system of this kind seems to me impossible. Therefore I forbear to discuss most of Tolman's systematic elaborations, as his to me most unsatisfactory chapters on emotions, appetites and aversions with their blindness for the real nature of the esthetic side of human activity (279), and as his chapter on language. When he says: "Speech, in the last analysis, is but a 'high-faluting' 'tool,' not different in essence from other tools such as 'strings,' 'sticks,' 'boxes,' and the like" (236), he seems to fall a victim to the demands of his system. Personally I believe this proposition to be wrong, but I should require a vast and organized body of facts before I would make any positive statement one way or the other.

Space forbids me to discuss the highly suggestive chapters on learning (319–391). The interpretation of the conditioned reflex seems to me excellent, the discussion of trial and error learning, though not final, a very important contribution to the subject.

I add only a few words about the last part, "This System *Qua* System." I refrain from discussing Tolman's own ultimate systematic conclusions, following his injunction: "But may neither you nor we seek to hold up these propositions, save in a somewhat amused, a somewhat skeptical, and a wholly adventure-seeking and pragmatic behaviour-attitude" (394). I must only clear up some misunderstandings of gestalt theory. Tolman distinguishes three kinds of psychologies: "Individual psychology," chiefly concerned with individual differences, "normative psychologies" (of which gestalt psychology is a typical instance), whose main interest lies in "the 'normal mental processes' or 'events' which are supposed to go on in every 'normal' or 'average' individual, of the given species, in

response to such and such stimuli" (396), and the "complete psychologies," which are "concerned both with what specific standard stimuli do to many different individuals (individual psychology) and with what any given stimulus does to the average or standard individual (normative psychology)" (396–7). The reviewer must object to finding gestalt theory ranked as a normative psychology. Gestalt theory does not introduce the "normal" or "average" individual as the object of its system; rather does it try to establish general laws about the dynamics of behavior, which are as valid in the normal as in the abnormal individual, in the adult as in the child, in the human as in the subhuman organism. The law of gravitation has as its object not the solar or the galactic system, but *any* system of masses; the actual motions of these masses will, according to the same law, be different in different distributions. No real event is described by a law alone. All description requires "topographical" conditions as well. Before one knows the laws, one must try to keep the topographical conditions as constant as possible, and therefore most of the experimental work of the gestalt school was done with normal adults. But enough work has been done with abnormal persons, children, and animals to prove that the normal adult is not the object matter of gestalt psychology, and that the same laws, albeit leading to different results under different conditions, are found valid throughout. Therefore gestalt psychology is a complete psychology in a sense different from Tolman's. Its completeness or generality is not numerical, *i.e.*, applicability to as great a number of cases as possible, but a generality which knows no exception, and would still be a generality, if the same set of conditions did not occur more than once.

Secondly, Tolman's interpretation of the relation between sensation and gestalten in gestalt theory is wrong. Since, however, I don't wish to turn this review into an exposition of my theory, I omit this point. But a last point must be mentioned: "Finally, it must be noted that Gestalt Psychology, like Structuralism, seems to be relatively weak and silent on the question as to how the . . . gestalts finally go over into and cause behaviour" (399–400). I cannot understand this criticism. Not only is there Lewin's field theory, which Tolman finds similar to his own (179), but long before Köhler had formulated a theory of eye movements, whose great significance for behavior in general I had discussed in "The Growth of the Mind." In this theory forces in the perceptual field exert a direct controlling influence on muscle action. No other theory of behavior

is as concrete as this theory except the crude neural bond theory which Tolman himself rejects with the best of reasons. One cannot understand gestalt theory, unless one sees that it applies the laws of organization to both the receptor and the effector side of behavior, or rather to the whole of behavior of which perception and action are but two interdependent parts.

I should like to discuss Tolman's concept of capacity; indeed, I should like to discuss ever so many more problems treated in his book. But a review of a book must not become a book itself, and therefore it is meet to come to an end.

The reader of this book will find his interest held from the first to the last page. He will frequently be delighted by the ingenuity of experimental decisions and the cleverness with which the material is handled in an argument; he will also, probably, be disturbed by the complexity of the terminology, and sometimes, as I did, shake his head at what he may consider the absurdity of the standpoint. But always will he like the author, and at the end he will feel indebted to him for the wealth of suggestions he has imparted. No short review can do justice to this book, least of all a review like this, which, in order to be constructive, had to be so critical. Not every reader will disagree with the author as widely as the reviewer has done, for, many readers will applaud the concluding sentences of the book in which the reviewer seems to discern the keystone of the author's system: "But what outside reality may be, in and for itself, abstracted from all human behavioural needs and all human behavioural capacities, we do not, cannot, and need not know" (431).

K. KOFFKA.

Smith College.

ROBINSON, EDWARD S. *Association Theory To-day: an Essay in Systematic Psychology.* New York: The Century Co., 1932. Pp. viii+142.

This little book, written by the person best qualified to write it, is exactly what its title asserts: a systematic and analytic exposition of the generalization of association as it exists to-day. It is not a defense of associationism, except as historical continuity is always a defense of conservatism. So often a new word like *insight* (Robinson shares my view here) is only a new word. The new facts for which the new word was created can also be formulated under the old words. To speak of *Association—To-day* is, therefore, not to link incompatibles. Robinson makes that clear, whatever our verbal predilections.

Presumably the book must be read forward, but it is perhaps well to remember it in the reverse order of chapters. An excellent introduction is the short sixth, and last chapter. It gives the law of association, the big generality that all the other text is about. The two preceding chapters give the inductive ground for the generality, the ten laws of association, more or less. Finally the first three chapters deal with fundamental distinctions, the nature of the materials of association, the use of terms, and so on—the cod liver oil with all the good vitamins that one takes before the meal. (I think Robinson as a child saved his icing until last.)

The law of association—and let us be clear that Robinson does *not* use this term, for he speaks only of *laws*—the law of association is A=f,x,y,z, . . .), and where A is "associative strength" and x,y,z, . . . are "factors demonstrably related to associative strength." The book is about the laws of association, A=f(x), A=f(y), *etc.* The general formula, which *is* Robinson's, is a way of saying that the problem of association is a problem of the statement of sequential functional dependencies. A is the dependent variable, the effect, the instigated term; x, y, z are the independent variables, the causes, the instigators. This is all there is, Robinson notes casually, to the notion that association is atomistic. Any functional dependency must be analyzed for significance into two related terms, as A and x in A=f(x). (If this be atomism, let who will make the most of it.)

The account of the independent variables or instigators, x, y, z, is the account of the laws of association. What are the laws? (1) *Contiguity* certainly, the generally accepted, fundamental law of temporal togetherness. (2) *Assimilation* inevitably, the vicarious functioning of one instigator that is related to or similar to the primary instigator. (3) Also *assimilation* among the instigated processes, where a given instigator gives rise to a different instigated activity. (I say "good morning" to the stranger whom I met but last night.) (4) *Frequency,* the first mentioned and best known of the quantitative instigators. (5) *Intensity* of something or other, possibly sensory, possibly affective, possibly both, but something that results from the degree of the associated terms themselves, rather than of the relation between them. (6) *Duration* possibly, although is a little dubious and needs more study. (7) *Context,* the meaningful ground in which the instigator is set, the congruous apperceptive systems that belong to the instigator. (8) *Acquaintance,* the associative advantage of familiar terms over unfamiliar, of sense

over nonsense. (9) Possibly *composition,* the intrinsic susceptibility of some psychological activities for association as compared with others. (10) And certainly *individual differences,* as everyone knows. Robinson argues all of these matters out at greater or less length for the more important part of the book.

These laws are taken at a relatively empirical level. Each of them is, usually, what certain experiments show. New ones may be added; old ones discarded; combinations effected. Even those that are conjectural are important, because they lead to experimentation and are thus determined. Systematization goes hand in hand with experimentation, and Robinson believes that his book is not only a summary of principles learned but also a suggestion of research to be done. All right; he was bound to take just that ground in modernizing the concept of association, in showing that association is what the experiments on association are about.

Nevertheless, I am not quite happy about this hodge-podge of laws. Robinson says that it is not a hodge-podge: $A=f(x,y,z\ldots)$ is the prince of unities, like Area=f(length, breadth). He is quite right, I think, about the formulae but not, I think, about association. The difficulty arises, I venture to suggest, because he does not define A=associative strength. Fifty-nine careful, even hair-splitting pages about x, y, z, and not even a pair of pages about A! One might reasonably ask for twenty. If we are being empirical and not speculative, we must keep to the experiments, and there we have all the different measures of associative strength, $A=f(x)$, $B=f(y)$, $C=f(z)$, ..., a system that is susceptible of being a hodge-podge. Well, we all must choose, and Robinson I think has chosen empiricism, safety, accuracy and multiplicity of laws: the maximum of what science begins with and thus also the maximum of what science tries to obliterate. However, if Robinson could have defined A as well as x, y, z in $A=f(x,y,z, \ldots)$ then we should have unity. We should have all the instigators related in terms of the single instigated, and we could surprise him by answering the question, "How many times is how contiguous?" For myself I think we shall not look very proudly at the psychology of association that shirks the answer to such a question.

Most of the argumentation, most of the book, seems to me to be inevitably consequential: the logic holds. I am not always so certain, however, that it is important. Why should we make and labor the distinction between quality and quantity, especially if we are to keep at the empirical level? And why do we need a whole chapter out

of our precious six on the distinction between *simultaneous* and *successive?* Let me rewrite this chapter for Robinson here and now in one long sentence: Association is a relation between a prior instigator and a subsequent instigated term; hence it appears normally to be successive; however, association is very complex, and a complete instigator may almost instantaneously give rise to a complete instigated resultant and the instigator and its resultant exist together in a single organization, so that simultaneity may sometimes seem as much the rule as successivity.

Robinson has a dualism, although it is not Cartesian. A movement can be an instigated term in an association, but not an instigator. A stimulus can not enter into an association at all, although a stimulation may be an instigator. The great wall of psychology is between the "internal categories" and the "external categories," and movement may be inside although a stimulus is always outside. The initial difficulty here is hardly more than terminological. It might be said of the *stimulus-response* relationship that each term intends the other, that neither is self-contained, and that the stimulus is not a stimulus except as it is a stimulus to a response, and conversely. It is probably such a stimulus that Robinson calls *stimulation* and such a reponse that he calls *movement*. However, this is not quite all of the difficulty. Robinson is explicit in his doctrine of the stimulus, and he settles the stimulus firmly outside of psychology, shutting the door in its face. (See pp. 29–31, 38, 44 f.) Why can not a stimulus be associated?

It can not be associated because (for Robinson) it is just an object, a thing, in outer physical darkness. It must have on a "wedding garment" before it can even be introduced to an instigatable activity. The implication seems to be that all those at the associative marriage feast are in intimate or direct psychological contact. Actually there is no such intimacy at all. A retinal stimulation presumably can become the instigator of an adrenal secretion, and we can guess something as to the large number of intermediate terms there must be for those two remote activities to get introduced to each other so that the association can be established. Similarly a snake in the grass may be the "instigator" of a retinal stimulation (if I ignore Robinson's taboo and use this word in its general meaning). What good can come from interpolating in the casual series from physical process through afferent and efferent activities into another physical process a threshold with the stimulus on one side

and stimulation on the other? If the adrenal process is instigated by any one of a hundred sequential events that come after the retinal stimulation, why not also by a few events in the series just prior to this point? And then, of course, a movement by way of proprioception could become an instigator in a circular association.

All in all this book is timely. Robinson has carefully and conscientiously altered the associative garment so that it fits the growing child, Learning—really an excellent fit. Just the same it is not yet a beautiful coat. I think what I want is to have Robinson dye it a single color, so that it will look less patchy.

EDWIN G. BORING.

Harvard University.

LINDWORSKY, JOHANNES. *Theoretical Psychology.* (Tr. from the German by Harry R. DeSilva.) St. Louis, Mo.: B. Herder Book Co., 1932. Pp. xi+145. $1.25.

Contemporary American writing upon the more general theoretical problems of psychology has, if one excepts the attention given to *gestalt-psychologie,* been little concerned with current European theorizing. It is the more fortunate, therefore, that Professor DeSilva has made readily available, by his excellent translation, an interpretation of the field written from a different background and occupied with a somewhat different set of facts. Lindworsky has had, as his translator points out, "a careful training in experimental technique under the late Professor Oswald Külpe, and an incomparable opportunity to obtain a view of the whole field under that most scholarly of psychological encyclopedists, Professor Joseph Fröbes" (p. iii). He has, since 1909, been engaged upon the problems of thought and volition, with occasional excursions into related fields, including the *Experimental Psychology,* also translated by DeSilva. The present book, while by no means dominated by its author's own special research interests, gives evidence of their definite influence.

The common subject matter of both experimental and theoretical psychology is, for Lindworsky, conscious phenomena and their connections. Psychology has the peculiarity that it is only indirectly accessible to quantitative measurement and that, being youthful, it can scarcely hope to provide generalizations which are comparable in range to those of physics, but must be content with more modest formulations. The method of theoretical psychology is "to recognize the like in the different, to perceive what is common to all and to

formulate it into laws . . ." (pp. 6–7), in an effort to arrive at a systematic panorama. This leads him to organize the facts under the two major headings, "The Laws of Content" and "The Laws of Process."

A description of the laws of content leads him at once to the materials of content and thence to the elementary experiences, which are six in number: (1) the "I-characteristic," which pervades the knowing of experience; (2) sensations; (3) sensory feelings; (4) the knowing of facts; (5) the processes of recognition; and (6) the striving attitudes. These are, in turn, described, but it is noteworthy that, although introspection of one kind is implied throughout, the author feels no necessity either to explicate or to defend it. This is interesting testimony to the particular perspective in which he sees his field. To psychologists who have lived through the period of militant objectivism, it will seem a strange oversight. After the laws of the materials come those of the form or constitution of content in which are expressed the coördination in which the different experiences actually appear. The description and logical ordering of the contents are carefully done and contain much that is unique, but particular attention should be called to his theory of relations, which is intrinsic to a great deal of the discussion. The relational contents are not elementary, but are derived from the true elements by a "combination of the reflex experience with the reproductive processes"[1] (p. 122).

The laws of process state the conditions under which mental phenomena occur, and discussion begins with the regular sequence of any two such phenomena. To a sequence of the type illustrated by successive contrast, three main attributes attach: indifference of content, which means that the sequence may not be understood from the contents themselves, preëminence of parallel physiological events; and palpability (integration through palpable sensory contents and sensory feelings). Here no purely psychical laws are possible, but in the second kind of experience, illustrated by apprehension of relation and by striving, a physiological parallel is lacking, the sequence is to be understood from its contents, the processes are impalpable, and here "*real psychical laws* are possible" (p. 54).

[1] The reflex experience is one of the recognizing attitudes of the I and is an expression borrowed from scholasticism to refer to self-conscious experiencing. See, also, J. Lindworsky, Revision einer Relationstheorie, *Arch. f. d. ges. Psychol.*, 1924, **48**, 248–289.

The laws of psychophysically conditioned phenomena and sequences are, as one would expect from this distinction, separately treated. The ultimate law of the former is: *"The sequences of phenomena are conditioned by psychophysical processes. These take place as if bodies which live (i.e., mature, develop, age) and are equipped with a transposable, limited psychophysical energy, vibrated in an elastic medium"* (p. 96). The resonance hypothesis implied in this law is given somewhat extended treatment and made to bear a considerable weight of fact. The laws of the purely psychically conditioned sequences are those of recognition and striving. In the course of the development of both sets of laws there is much interesting logic and much important criticism of the theories of others, particularly of G. E. Müller and of Koffka.

The book is a concise and closely reasoned attempt to weave experimental fact into a theoretical system which must, in the end, by virtue of its premises, transcend experiment in a particular direction. It suggests the writing of the German and British systematists of the last century, although Lindworsky's acquaintance with experimental fact is wider by far than theirs and his logic keeps closer to the laboratory. But a theory tinged with Scholasticism must finally forsake the laboratory with sharp renunciation, and the appeal to experimental facts is, as his definition of subject matter would assure one, to the facts of conscious phenomena and their sequences, with, in the case of part of them, the implied physiological substrate. The quantitative behavior data of contemporary research, especially outside of Germany, play no part. The theorizing, likewise, shows little touch with the now somewhat declining wars of American systems, nor does it take cognizance of many of the newer contemporary movements in Germany. Within the domains of its intent, however, it is well-knit and systematic, without becoming closed, formal, and inflexible. It aims at organization without militant defense of theory, at calm systematizing of particular classes of facts with no blare of trumpets and no proclamation that the world psychological is about to be saved thereby.

Psychologists could be found, no doubt, who would decry such a book as valueless, if not pernicious, by reason of its concern with conscious life alone and its scholastic tenor. One, however, can disagree sharply, as this reviewer does, with the author's definition of the field, with many of his special theories, and with his whole systematic *Weltanschauung,* and yet read the book with zest and with

respect. The domains of psychology are too diversely luxuriant to be confined to any single system yet formulated, but serious attempts to organize the field are valuable. Systems, moreover, with which one fails to agree are needed correctives, and Lindworsky's emphasis upon a point of view not now highly regarded in this country is interesting both as a phenomenon in its own right and as systematic counterpoise.

JOHN A. MCGEOCH.

University of Missouri.

A REPLY TO MR. KOFFKA

Koffka's review shows what is wrong with my[1] book from a gestalt point of view. For that reason I find it both acceptable and exciting. So much so, that I have pleaded to be allowed to have this answer included in the same issue.

I select four of Mr. Koffka's difficulties to reply to. They are: (1) my identification of the sufficient criterion of a mental process, with the definition of that mental process; (2) my implication that a mental field is not as such private—that in short, others can know my mental field as well as I can myself; (3) my assumption of no *necessary* structural, *i.e.*, constitutive, correspondence between mental processes, defined behaviorally, and the brain activities underlying such mental processes; and (4) my assertion that the sign-gestalt is the pattern for all gestalts. Let me consider these four one at a time.

(1) The identification of the sufficient criterion of a mental process with the definition of that mental process. I shall still assert that an identification of sufficient criterion with the definition of a concept is scientifically legitimate. I believe that it is all that any good science does. I believe, in short, contrary to Koffka, that the "excursion of a pointer on a seismograph" (or rather the excursion of all the pointers on all the seismographs, together with all the other normal "behavioral" results of an earthquake) do "define" the concept—earthquake—and that the actual occurrence of such pointer readings and of such other normal accompaniments on any particular occasion denote an actual occurrence of an actual earthquake on that occasion. Or, in other words, I am merely contending for that general point of view which has come to be known as "operationalism."[2]

[1] To avoid clumsiness the first person singular has been used throughout this reply in spite of the double authorship. All the statements are the conjoint product of both authors.

[2] Briefly, operationalism is the view that when we use concepts, whether they be psychological concepts such as "cognition" or physical concepts such as "length," we *mean* nothing other than, nor more than, the "operations"—the sorts of things we do—to verify the propositions in which these concepts occur. A set of operations thus both "defines" and supplies "the sufficient criteria" for the explication of the meaning of a concept. See P. W. Bridgman, *The Logic of Modern Physics*, Chapter I. See also George Boas and Albert E. Blumberg, "Some Remarks in Defense of the Operational Theory of

The task of psychology as I now see it is to discover the various "forms" which the equation which connects behavior to stimuli and to the other independent variables may take. That is, psychology, aiming as it does at the prediction and control of behavior, seeks to write the various forms of the function f_e in the equation:

$$R=f_e\ (S, H, T, P)$$

in which, R is the response,
S is the total stimulus field at the moment,
H is the heredity of the organism,
T is its specific past training,
P is the physiological drives of the moment.

It is to be noted that the form of this function, f_e, is different in different individuals and changes in the same individual from moment to moment. That is, changes in the independent variables S, H, T and P will not only produce different values of R, when f_e remains constant, but will also tend (not "mathematically" but "physiologically") to induce different values in f_e itself. And it is these various different values of f_e which I would define as the mental processes or states. When one mental process or state is in force one set of R's result from given S's and P's, H's and T's. When another mental state or process is in force another set of R's will result. And we know and *define* such types of mental processes or states only by these effects in thus causing such and such responses from such and such stimuli, physiological conditions, heredity and past training.[3]

Actually, however, it is going to be difficult to write these defining forms of the function f_e in algebraic or equational terms, the relations are so involved and the very variables themselves are so

Meaning," *J. Phil.*, 1931, **28**, 544–550; and Herbert Feigl and Albert Blumberg, "Logical Positivism," *J. Phil.*, 1931, **28**, 281–296.

[3] See E. C. Tolman and I. Krechevsky, "Means-End-Readiness and Hypothesis—A Contribution to Comparative Psychology." *Psychol. Rev.*, 1933, **40**, 60–70, esp. pp. 62 f. And I would also refer the reader in this connection to a forthcoming article by Jack Buel on "Differential Errors in Animal Mazes" for a further detailed illustration of this point of view applied to the specific problem of maze learning.

At the time of writing the book I called the mental states and processes, not the form of the total equation, but "intervening variables" in that I then broke down the total equation into two sub-equations of the types $R=f_1(M)$ and $M=f_2(S, H, T, P)$ in which M designated the "intervening" mental states or processes. It seems, however, that it is more fruitful, although mathematically not essentially different, to adhere to the one total equation and to talk not of M's but of the "forms" which the total function f_e assumes from moment to moment.

mutually interacting. A more hopeful attack at the present time seems to me therefore that of translating these varying forms of the function into diagrammatic terms (on an analogy with the procedure of analytical geometry). If we can not at present write equations, we *can* draw diagrams. And these will serve as a substitute for the equations. We can draw pictures of "environmental fields" (or, as I would prefer to call them, means-end-fields), as these exist for the behavior of the given organism at given moments. It is this that Lawin does. And his pictures of "environmental fields" with their topological structures, their valences, their barriers, etc., are really, as I see it, attempts at writing the momentary forms of the function f_o in the equation $R=f_o$ (S, H, T, P) diagrammatically. From such diagrams, if they be correct, one can predict what R's are going to follow from given S's. Or *pari passu* one can, by pointing to the diagram, give a definition of the given mental states and processes of the given moment. The diagrammatized field gives a picture of the function f_o which holds at the moment. This momentary "function," or this field, whichever you prefer, is, however, only an inferential state of affairs. It has to be both discovered and defined by virtue of its effects. It is a potentiality for such and such a totality of possible effects and this is all that, in the last analysis, one really can say about it.

The fact that the organism under consideration may be I myself and that, if it be so, then I can save time for you, the observing psychologists, by "introspecting" for you, does not in any way change the essence of the situation. The nature and order of your final information, if I introspect, is exactly the same as it is, if, instead, I grossly run about in a maze or choose between the colors in a discrimination-box. In either case you learn simply that I am suffering from a mental process—*i.e.*, one whereby I then and there would tend to respond to such and such stimuli in such and such ways. This tendency or this mental process you may try to write algebraically by stating the function f_o which holds in me at the moment, or, you may depict it in terms of the "means-end-field" holding for me at that moment. But neither equational function nor diagram will be a statement of my "direct experience." Direct experience does not enter into our science—into psychology no more than into physics.[4] Both equation and field are, if you will, as Koffka

[4] *Cf.* E. G. Boring, *The Physical Dimensions of Consciousness* (New York: The Century Co., 1933), Chapter I. Boring states a view akin to the one here advanced concerning the subject-matter of science. Paraphrased, his position

suggests, "fictions" or perhaps better constructs which are not directly observable even by the "subject" himself. But they are fictions or constructs which are subject to test and revision in terms of their pragmatic usefulness. They are constructs in the very same sense in which atoms and electrons are constructs. And it is always with constructs in this sense that any science deals.[8]

All my knowledge of expectations, cognitions, and the like, was built up by observing either my own or other people's behavior. I didn't "introspect" to find out what an "expectation" or a "cognition" or a "learning" was. In the book I merely tried to make explicit the behavioral criteria and assumptions which we all make and I attempted to demonstrate such criteria and assumptions experimentally. By such experimental demonstrations I did not attempt to assert that an expectation is the same as any of its effects—but rather that it was the same as the potentiality of *all* of its effects.

(2) The implication that a mental field is not as such private—that, in short, others can know my mental field as well as I can myself. Mr. Koffka asserts the contrary. He insists that consciousness is private. He keeps pointing out that "after all" we begin

is that science, psychology as well as physics, begins not with "direct experience" but with tentative hypotheses, *i.e.*, with tentatively ordered concepts, and ends, not in "direct experience," but by discovering relationships between "constructs," *i.e.*, by discovering natural laws. Its purpose is to describe, predict and control phenomena—psychological as well as physical, not to reproduce or restimulate "direct experience." Such restimulation is the function of poetry. Boring's trenchant criticism of behaviorists who "ignore" consciousness does not apply to molar behaviorism, which rather restates and makes explicit in behavioral terms the "conscious field" or "structure"—leaving, of course, the *stuff* of consciousness, as physics leaves the stuff of matter, just where it is.

[8] Properly speaking, a construct is a formulation in terms of "operations," "formulae," etc., of what may naïvely be thought of as an "entity" or "thing." When it is said that an electron is a "construct" what is meant is that all that we really know of an electron can be stated completely by the formulae and operations of the physicist—no little round balls remain outside. Similarly all that we really know of a "mental process" can be stated completely by the formulae and operations of the stimulus-response connections in which this mental process occurs. This does not mean that electrons and mental processes do not exist; it merely means that they "occur" only as variables and constants in physical and behavioral equations.

The point of view here taken derives not only from P. W. Bridgman's *The Logic of Modern Physics*, but also from C. I. Lewis' *Mind and the World-Order* and from the school of contemporary philosophers variously known as Logical-Positivists, Neo-Positivists and Physicalists. (The doctrines of the latter are to be found in the German periodical "Erkenntnis.")

with our "own conscious fields." But this is misleading. What we begin with is not private fields but fields-to-which-we-respond. And there is nothing private about the latter. The fields to which we respond are as I have just been trying to show objectively discoverable functions. "Raw feels," and "direct experience" (if there be such) may be private—that is, non-communicable[6] but "fields" are communicable—both mine and those of the gestalt psychologists. Otherwise neither they nor I would have attempted to write books about them. My description of the rat's field was not in terms of *my* conscious (*i.e.*, supposedly private) field but rather in terms of *his* (the rat's) behavior.

(3) The assumption of no necessary structural, *i.e.*, constitutive, correspondence between mental processes, defined behaviorally, and the brain activities underlying such mental processes. Mr. Koffka is distressed because I have called molar behavior an emergent phenomenon and have asserted that it was not the same thing and *need not* have the same properties as the brain physiology underlying it. My original statement on this point had really two sub-propositions involved in it.

The first of these was a compound one to the effect (a) that the study of behavior *qua* behavior (*i.e.*, the study of psychology) is not the study of physiology, although (b) all the laws and functions of behavior may ultimately rest upon those of physiology. The equations that we write (the diagrammatic fields that we draw) in the study of behavior (*i.e.*, in psychology) are not those which we write and draw in brain physiology. Psychology deals with a different set of variables from those dealt with by brain physiology. Psychology concerns itself with valences, expectations, discriminanda, sign-gestalts. Brain physiology deals with neural patterns and activities. The former are presumably dependent upon the latter. But I see no reason to identify them with the latter nor even to accept the principle of *isomorphism*, viz., the principle that the dimensions of the two series must be exactly alike. They need be no more alike than the "dimensions" of the gases H and O are like those of the liquid, water. The transformation equations connecting the one series with the other will have to be empirically discovered and I do not believe that these equations necessarily will be simple ones.

My second sub-proposition seems to have been that brain physiology is molecular (as contrasted with the molar-ness of behavior *qua* behavior). This, if implied, was a slip-up. The evidence is undoubt-

[6] *Cf.* D. Rynin, The Nature of Communication, *J. Philos.*, 1932, **29**, 505–515.

edly piling up that brain physiology as well as behavior is molar-istic or "gestalt-y." But again I assert that this does not necessarily mean the further principle that all the dimensions and properties of the behavior series are either identical with or even very much like those of the underlying brain physiology series. The two series are, I said, in one-to-one correspondence but I meant this only in the sense of a functional correspondence (one without loop-holes where a soul could suddenly "butt in"). But the nature of the transformation equations between the two series will have to be worked out. They cannot, I believe, be known beforehand in terms of any nice, simple *a priori* principle such as that of isomorphism.[7] Knowing all there is to know about brain physiology will not in itself tell me anything about valences, expectations, etc., until I have learned the nature of the correlating equations; and, *vice versa,* knowing about valences, expectations and discriminanda will not tell me the nature of the underlying physiology. To this extent, and *to this extent only,* I am an emergentist. Each science has to discover its own facts and laws. And then both have to join together to discover the nature of the transformation equations which will convert the one into the other.

For Koffka there are two series—a brain physiology series which is molar and a psychic series which is also molar. And this latter series is, for him, in *isomorphic* correspondence with the former. For me there are also two series. There is a brain-physiology series, which I am also perfectly willing to admit may be largely or wholly molar in character, and there is a behavior series (also molar) which is predictable from the former—but predictable *only after* one has *empirically* discovered the nature of the transformation equations[8] which connect the two. I am skeptical of any single convenient, *a priori,* equation such as that of *isomorphism.*

(4) The assertion that the sign-gestalt is the pattern for all gestalts. Here I admit that I was on doubtful and tentative ground. As I now see it, the only real virtue in stressing sign-gestalts as against "pure" gestalts is to emphasize that pure (*i.e.,* non-sign-) gestalts, if there be such, are not, in scientific discourse, direct experiences but conceptual entities, *i.e.,* statements of certain types of relationships or "organization," and further that these pure gestalts always fall into larger means-end (*i.e.,* sign-significate) fields. It is not so much the fact of, or reality of, or importance of, sign-gestalts

[7] And the premature acceptance of such a principle may well militate against the following up of fertile leads and may well tend to warp the theoretical interpretations in both sciences.

[8] Analogous, perhaps, to the Lorentz transformation equations in physics.

which I especially want to emphasize as it is the fact that all gestalts are shot through with means-end relationships and, finally, that all gestalts (as I especially emphasized in the case of the sign-gestalt) should be opened to genetic explanation in terms of the organism's history (ontogenetic or phylogenetic). I am now perfectly willing to admit that the sign-significate relation (definitive of the sign-gestalt) may be only one of the important types of means-end-relations; round-aboutness, obstructions and barriers would seem to be equally important. The points to be emphasized are that gestalts are conceptual, that they are determined by the genetic history of the organism and that they are shot through with means-end-relations.

To conclude: Even if I have not been able to answer even these few of Mr. Koffka's criticisms satisfactorily, I can but thank him for the trenchancy and the fundamentalness of his attack.

EDWARD C. TOLMAN.
JOHN HOROWITZ.

University of California.

SKIN POTENTIALS

Broxon and Muenzinger (1) are quite correct in calling attention to my misinterpretation (2) of their experiment (3). Their demonstration of the fact that there is a skin response shown as a change of potential in a circuit allowing no current flow is the first of its kind. There was, of course, strongly presumptive evidence that this was true. Darrow (4) and Jefferess (5) had both shown that there was a change in skin potential when isolated from skin resistance in circuits using current flow. Boyd (6) has measured static potential levels from the body.

With respect to the designation of the quadrant electrometer as an old-fashioned instrument, I use the descriptive adjective "old-fashioned" in the usual sense of the word, *i.e.*, one which has for the most part been out-moded by newer developments. In many physical laboratories to-day the quadrant electrometer has been replaced by the General Electric Company FP54 Pliotron tube which has a grid current of 10^{-15} amperes and an input impedance of 10^{16} ohms. In still other experiments the circuits of Fosbinder and Schoonover (7), Dubois (8) and others are frequently used to measure potentials. In the study mentioned above, Boyd used a vacuum tube set-up.

edly piling up that brain physiology as well as behavior is molar-istic or "gestalt-y." But again I assert that this does not necessarily mean the further principle that all the dimensions and properties of the behavior series are either identical with or even very much like those of the underlying brain physiology series. The two series are, I said, in one-to-one correspondence but I meant this only in the sense of a functional correspondence (one without loop-holes where a soul could suddenly "butt in"). But the nature of the transformation equations between the two series will have to be worked out. They cannot, I believe, be known beforehand in terms of any nice, simple *a priori* principle such as that of isomorphism.[7] Knowing all there is to know about brain physiology will not in itself tell me anything about valences, expectations, etc., until I have learned the nature of the correlating equations; and, *vice versa,* knowing about valences, expectations and discriminanda will not tell me the nature of the underlying physiology. To this extent, and *to this extent only,* I am an emergentist. Each science has to discover its own facts and laws. And then both have to join together to discover the nature of the transformation equations which will convert the one into the other.

For Koffka there are two series—a brain physiology series which is molar and a psychic series which is also molar. And this latter series is, for him, in *isomorphic* correspondence with the former. For me there are also two series. There is a brain-physiology series, which I am also perfectly willing to admit may be largely or wholly molar in character, and there is a behavior series (also molar) which is predictable from the former—but predictable *only after* one has *empirically* discovered the nature of the transformation equations[8] which connect the two. I am skeptical of any single convenient, *a priori,* equation such as that of *isomorphism.*

(4) The assertion that the sign-gestalt is the pattern for all gestalts. Here I admit that I was on doubtful and tentative ground. As I now see it, the only real virtue in stressing sign-gestalts as against "pure" gestalts is to emphasize that pure (*i.e.,* non-sign-) gestalts, if there be such, are not, in scientific discourse, direct experiences but conceptual entities, *i.e.,* statements of certain types of relationships or "organization," and further that these pure gestalts always fall into larger means-end (*i.e.,* sign-significate) fields. It is not so much the fact of, or reality of, or importance of, sign-gestalts

[7] And the premature acceptance of such a principle may well militate against the following up of fertile leads and may well tend to warp the theoretical interpretations in both sciences.

[8] Analogous, perhaps, to the Lorentz transformation equations in physics.

which I especially want to emphasize as it is the fact that all gestalts are shot through with means-end relationships and, finally, that all gestalts (as I especially emphasized in the case of the sign-gestalt) should be opened to genetic explanation in terms of the organism's history (ontogenetic or phylogenetic). I am now perfectly willing to admit that the sign-significate relation (definitive of the sign-gestalt) may be only one of the important types of means-end-relations; round-aboutness, obstructions and barriers would seem to be equally important. The points to be emphasized are that gestalts are conceptual, that they are determined by the genetic history of the organism and that they are shot through with means-end-relations.

To conclude: Even if I have not been able to answer even these few of Mr. Koffka's criticisms satisfactorily, I can but thank him for the trenchancy and the fundamentalness of his attack.

EDWARD C. TOLMAN.
JOHN HOROWITZ.

University of California.

SKIN POTENTIALS

Broxon and Muenzinger (1) are quite correct in calling attention to my misinterpretation (2) of their experiment (3). Their demonstration of the fact that there is a skin response shown as a change of potential in a circuit allowing no current flow is the first of its kind. There was, of course, strongly presumptive evidence that this was true. Darrow (4) and Jefferess (5) had both shown that there was a change in skin potential when isolated from skin resistance in circuits using current flow. Boyd (6) has measured static potential levels from the body.

With respect to the designation of the quadrant electrometer as an old-fashioned instrument, I use the descriptive adjective "old-fashioned" in the usual sense of the word, *i.e.*, one which has for the most part been out-moded by newer developments. In many physical laboratories to-day the quadrant electrometer has been replaced by the General Electric Company FP54 Pliotron tube which has a grid current of 10^{-15} amperes and an input impedance of 10^{15} ohms. In still other experiments the circuits of Fosbinder and Schoonover (7), Dubois (8) and others are frequently used to measure potentials. In the study mentioned above, Boyd used a vacuum tube set-up.

The apparent resistance change obtained by means of the Hathaway apparatus, with which Broxon and Muenzinger correlated their change in electrostatic E.M.F. would be affected by a recent study by Forbes (9) which shows that the Hathaway deflections are affected by gross body impedance level. Whether correction for this factor would increase or decrease the correlation cannot be predicted off-hand.

The essential point which Broxon and Muenzinger have shown, viz., that there is a pure E.M.F. *response* of the skin is a unique demonstration and one which I erroneously interpreted in my review.

CARNEY LANDIS.

N. Y. Psychiatric Institute.

REFERENCES

1. BROXON, J. W., and MUENZINGER, K. F., Isolating the Factor of Skin Potentials. *Psychol. Bull.*, 1933, **30**, 256–257.
2. LANDIS, C., Electrical Phenomena of the Skin. *Psychol. Bull.*, 1932, **29**, 693–752.
3. BROXON, J. W., and MUENZINGER, K. F., Changes in Skin Potentials During the Psychogalvanic Reflex. *J. Genl. Psychol.*, 1931, **5**, 94–98.
4. DARROW, C. W., and HEATH, L. L., *Studies in the Dynamics of Behavior.* Univer. Chicago Press, Chicago, 1932, 59–327.
5. JEFFERESS, L. A., Galvanic Phenomena of the Skin. *J. Exper. Psychol.*, 1928, **11**, 130–144.
6. BOYD, W. D., The Electric Field of the Human Body. *Brit. J. Radiol.*, 1930, **3**, 128–135.
7. FOSBINDER, R. J., and SCHOONOVER, J., An Improved Method of Measuring Glass Electrode Potentials. *J. Biol. Chem.*, 1930, **88**, 605–614.
8. DUBOIS, D., A Vacuum-Tube Potentiometer Applicable for Use with Glass Electrodes of High Resistance. *J. Biol. Chem.*, 1930, **88**, 729.
9. FORBES, T. W., The Hathaway Galvanic Apparatus as an Instrument of Research. (In press.) *Amer. J. Psychol.*, 1933, **45**.